

us-10-030-735-1.rag

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 17:59:29 ; Search time 68.4493 Seconds  
(without alignments)  
70.173 Million cell updates

/sec

Title: US-10-030-735-1  
Perfect score: 108  
Sequence: 1 KRFKQDGGWSHWSPWSS 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being pri  
nted,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result                      %  
                            Query

No.	Score	Match	Length	DB	ID	Description
-----						
1	108	100.0	17	2	AAR32449	Aar32449 Hu
man thr						
2	108	100.0	17	2	AAR55067	Aar55067 Fi
bronect						
3	108	100.0	17	2	AAR69776	Aar69776 Th
rombosp						
4	108	100.0	17	3	AAB02656	Aab02656 Hu
man thr						
5	108	100.0	17	4	AAB35336	Aab35336 Al
pha3bet						
6	108	100.0	17	6	ABU13826	Abu13826 Th
rombosp						
7	108	100.0	18	2	AAR79367	Aar79367 Th
rombosp						
8	108	100.0	18	3	AAB02667	Aab02667 Hu
man thr						
9	108	100.0	18	6	ABU13856	Abu13856 Th
rombosp						
10	108	100.0	18	6	ABU79104	Abu79104 Th
rombosp						
11	108	100.0	76	5	AAU74793	Aau74793 Hu
man thr						
12	108	100.0	183	5	AAU74791	Aau74791 Hu
man thr						
13	108	100.0	218	2	AAW40287	Aaw40287 Hu
man TSP						
14	108	100.0	218	2	AAU06182	Aay06182 Th
rombosp						
15	108	100.0	239	2	AAR40823	Aar40823 Hu
man thr						
16	108	100.0	441	2	AAW40288	Aaw40288 Hu
man con						
17	108	100.0	441	2	AAU06183	Aay06183 Th
rombosp						
18	108	100.0	546	4	AAU02915	Aau02915 An
giotens						
19	108	100.0	555	4	AAU02914	Aau02914 An
giotens						
20	108	100.0	731	4	AAU02913	Aau02913 An
giotens						
21	108	100.0	1152	3	AAB00042	Aab00042 Hu
man thr						
22	108	100.0	1152	5	AAU74771	Aau74771 Hu
man thr						
23	108	100.0	1152	5	ABB82285	Abb82285 Hu



us-10-030-735-1.rag

man thr	24	108	100.0	1170	4	AAB74450	Aab74450	Hu
man var	25	108	100.0	1170	4	AAB90800	Aab90800	Hu
man she	26	108	100.0	1170	5	AAE25030	Aae25030	Hu
man thr	27	108	100.0	1170	5	AAU75315	Aau75315	Hu
man thr	28	108	100.0	1170	6	ABP96780	Abp96780	Hu
man COP	29	108	100.0	1170	6	ABU03474	Abu03474	An
giogene	30	108	100.0	1170	6	ABG74673	Abg74673	Hu
man THB	31	108	100.0	1170	6	AAE36228	Aae36228	Hu
man THB	32	108	100.0	1170	7	ABR62059	Abr62059	Hu
man thr	33	103	95.4	18	3	AAB02686	Aab02686	Hu
man thr	34	103	95.4	76	5	AAU74794	Aau74794	Hu
man thr	35	100	92.6	18	3	AAB02669	Aab02669	Hu
man thr	36	94	87.0	18	3	AAB02671	Aab02671	Hu
man thr	37	94	87.0	18	3	AAB02670	Aab02670	Hu
man thr	38	94	87.0	183	5	AAU74789	Aau74789	Hu
man thr	39	94	87.0	300	3	AAB00041	Aab00041	Hu
man COM	40	94	87.0	1172	3	AAB19677	Aab19677	Hu
man thr	41	94	87.0	1172	3	AAB00043	Aab00043	Hu
man thr	42	94	87.0	1172	5	AAU76902	Aau76902	Hu
man Thr	43	94	87.0	1172	5	AAU74788	Aau74788	Hu
man thr	44	94	87.0	1172	5	AAE25031	Aae25031	Hu
man thr	45	94	87.0	1172	5	ABP68622	Abp68622	Hu
man pan	46	94	87.0	1172	6	ABG72839	Abg72839	Hu
man ang	47	94	87.0	1172	6	ABG74674	Abg74674	Hu

## us-10-030-735-1.rag

man THB	48	94	87.0	1172	6	AAE36411	Aae36411	Hu
man THB	49	94	87.0	1172	7	ADD31095	Add31095	Hu
man thr	50	92	85.2	15	3	AAB02685	Aab02685	Hu
man thr	51	89	82.4	14	2	AAR55065	Aar55065	Fi
bronect	52	89	82.4	14	2	AAR69775	Aar69775	Th
rombosp	53	89	82.4	14	6	ABU13825	Abu13825	Th
rombosp	54	89	82.4	183	5	AAU74796	Aau74796	Mo
use thr	55	89	82.4	1172	5	AAU74786	Aau74786	Mo
use thr	56	89	82.4	1172	5	ABB72334	Abb72334	Ra
t prote	57	87	80.6	71	5	AAU74792	Aau74792	Hu
man thr	58	86	79.6	18	3	AAB02679	Aab02679	Hu
man thr	59	86	79.6	18	6	ABU79105	Abu79105	Th
rombosp	60	85	78.7	18	3	AAB02678	Aab02678	Hu
man thr	61	82	75.9	57	2	AAy49504	Aay49504	Hu
man MET	62	82	75.9	57	2	AAy49507	Aay49507	Hu
man MET	63	82	75.9	57	4	AAG79062	Aag79062	Bi
ologica	64	82	75.9	57	4	AAG79061	Aag79061	Bi
ologica	65	82	75.9	57	4	AAB50006	Aab50006	TS
P1 doma	66	82	75.9	57	4	AAB50009	Aab50009	TS
P2 doma	67	82	75.9	59	5	AAE20783	Aae20783	Al
ternati	68	82	75.9	59	5	AAE20784	Aae20784	Al
ternati	69	82	75.9	60	2	AAR13641	Aar13641	Th
rombosp	70	82	75.9	60	2	AAW81489	Aaw81489	Th
rombosp	71	82	75.9	60	5	AAE20753	Aae20753	Th

## us-10-030-735-1.rag

rombosp	72	82	75.9	242	3	AAB00040	Aab00040	Hu
man COM	73	81	75.0	14	3	AAB02683	Aab02683	Hu
man thr	74	81	75.0	18	3	AAB02680	Aab02680	Hu
man thr	75	78	72.2	12	2	AAR69777	Aar69777	Th
rombosp	76	78	72.2	12	3	AAB02687	Aab02687	Hu
man thr	77	78	72.2	12	6	ABU13827	Abu13827	Th
rombosp	78	76	70.4	11	2	AAR32450	Aar32450	Hu
man thr	79	76	70.4	11	2	AAR55066	Aar55066	Fi
bronect	80	76	70.4	11	3	AAB02659	Aab02659	Hu
man thr	81	76	70.4	12	2	AAR79368	Aar79368	Th
rombosp	82	76	70.4	12	3	AAB02666	Aab02666	Hu
man thr	83	76	70.4	12	3	AAB02665	Aab02665	Hu
man thr	84	76	70.4	12	6	ABU13857	Abu13857	Th
rombosp	85	76	70.4	157	3	AAB08133	Aab08133	Am
ino aci	86	66	61.1	12	6	ABU13828	Abu13828	Th
rombosp	87	66	61.1	17	2	AAR69773	Aar69773	Th
rombosp	88	66	61.1	17	6	ABU13823	Abu13823	Th
rombosp	89	66	61.1	18	3	AAB02668	Aab02668	Hu
man thr	90	66	61.1	370	4	AAB50684	Aab50684	C.
elegan	91	65	60.2	1083	4	ABB61710	Abb61710	Dr
osophil	92	64	59.3	12	3	AAB02688	Aab02688	Hu
man thr	93	64	59.3	23	2	AAR13630	Aar13630	Th
rombosp	94	64	59.3	23	2	AAW81483	Aaw81483	Th
rombosp	95	64	59.3	23	5	AAE20782	Aae20782	Al

## us-10-030-735-1.rag

ternati							
96	64	59.3	961	5	AAU80188	Aau80188	Hu
man TSP							
97	64	59.3	966	5	AAU80189	Aau80189	Hu
man TSP							
98	62	57.4	11	2	AAR35392	Aar35392	Hu
man thr							
99	62	57.4	11	3	AAB02662	Aab02662	Hu
man thr							
100	61	56.5	12	2	AAR69778	Aar69778	Th
rombosp							
101	61	56.5	14	2	AAR69774	Aar69774	Th
rombosp							
102	61	56.5	14	6	ABU13824	Abu13824	Th
rombosp							
103	61	56.5	56	2	AAU49508	Aay49508	Hu
man MET							
104	61	56.5	56	4	AAB50010	Aab50010	TS
P2 doma							
105	60	55.6	84	6	ABR96137	Abr96137	Hu
man NOV							
106	60	55.6	730	6	ABU08955	Abu08955	Hu
man ADA							
107	60	55.6	997	4	AAB72283	Aab72283	Hu
man ADA							
108	60	55.6	1093	4	AAE02455	Aae02455	Mo
use sem							
109	60	55.6	1593	6	ABR96136	Abr96136	Hu
man NOV							
110	60	55.6	1686	4	AAE00934	Aae00934	Hu
man 278							
111	60	55.6	1686	4	AAE00913	Aae00913	Hu
man 278							
112	60	55.6	1686	4	AAB74944	Aab74944	Hu
man ADA							
113	60	55.6	1690	4	AAB86949	Aab86949	Hu
man met							
114	60	55.6	1784	3	AAB41379	Aab41379	Hu
man ORF							
115	60	55.6	4123	7	ABU62079	Abu62079	Hu
man jel							
116	60	55.6	4561	4	ABG30203	Abg30203	No
vel hum							
117	60	55.6	9222	4	ABG21064	Abg21064	No
vel hum							
118	59	54.6	142	4	ABB11408	Abb11408	Hu
man sec							
119	59	54.6	401	7	ADD18225	Add18225	Hu

## us-10-030-735-1.rag

man mol							
120	59	54.6	401	7	ADD18230	Add18230	Hu
man mol							
121	59	54.6	401	7	ADD18228	Add18228	Hu
man mol							
122	59	54.6	401	7	ADD18232	Add18232	Hu
man mol							
123	59	54.6	415	3	AAB58729	Aab58729	Br
east an							
124	59	54.6	558	4	ABB62610	Abb62610	Dr
osophil							
125	59	54.6	939	5	AAG68296	Aag68296	Hu
man sem							
126	59	54.6	954	5	AAG68295	Aag68295	Hu
man sem							
127	59	54.6	999	3	AAy94990	Aay94990	Hu
man sec							
128	59	54.6	1034	5	AAG68291	Aag68291	Hu
man sem							
129	59	54.6	1049	5	AAG68289	Aag68289	Hu
man sem							
130	59	54.6	1078	5	AAG68292	Aag68292	Hu
man sem							
131	59	54.6	1092	5	ABG34077	Abg34077	Hu
man Pro							
132	59	54.6	1092	6	ADA01364	Ada01364	Hu
man PRO							
133	59	54.6	1092	6	ADA43793	Ada43793	Hu
man sec							
134	59	54.6	1092	6	ADA43561	Ada43561	Hu
man sec							
135	59	54.6	1092	6	ADA01236	Ada01236	Hu
man PRO							
136	59	54.6	1092	7	ADA01120	Ada01120	Hu
man sec							
137	59	54.6	1092	7	ADA43677	Ada43677	Hu
man sec							
138	59	54.6	1092	7	ADA06939	Ada06939	Hu
man PRO							
139	59	54.6	1092	7	ADA08427	Ada08427	No
vel hum							
140	59	54.6	1092	7	ADB99720	Adb99720	Hu
man PRO							
141	59	54.6	1092	7	ADB87003	Adb87003	Hu
man PRO							
142	59	54.6	1092	7	ADB66158	Adb66158	Hu
man sec							
143	59	54.6	1092	7	ADB99836	Adb99836	Hu

## us-10-030-735-1.rag

man PRO							
144	59	54.6	1092	7	ADB99491	Adb99491	No
vel hum							
145	59	54.6	1092	7	ADB66042	Adb66042	Hu
man sec							
146	59	54.6	1092	7	ADC23440	Adc23440	Hu
man tra							
147	59	54.6	1092	7	ADC26133	Adc26133	Hu
man PRO							
148	59	54.6	1092	7	ADE04960	Ade04960	Hu
man PRO							
149	59	54.6	1092	7	ADE11266	Ade11266	Hu
man PRO							
150	59	54.6	1092	7	ADD88197	Add88197	Hu
man PRO							

## ALIGNMENTS

## RESULT 1

AAR32449

ID AAR32449 standard; peptide; 17 AA.

XX

AC AAR32449;

XX

DT 17-DEC-2001 (revised)

DT 10-JUN-1993 (first entry)

XX

DE Human thrombospondin repeat 2-derived peptide #19.

XX

KW thrombosis; coagulation; heparin binding inhibitor; type I repeat

.

XX

OS Synthetic.

XX

FH Key Location/Qualifiers

FT Region 1. .4

FT /note= "basic heparin-binding motif BBXB B= Arg,

Lys,

FT His, X= any AA"

FT Region 9. .12

FT /note= "preferred subsequence, lacks electrical c  
harge"

FT Region 12. .15

FT /note= "preferred subsequence, lacks electrical c  
harge"

XX

PN USN7801812-N.  
XX  
PD 15-DEC-1992.  
XX  
PF 06-DEC-1991; 91US-00801812.  
XX  
PR 06-DEC-1991; 91US-00801812.  
XX  
PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
XX  
PI Roberts DD;  
XX  
DR WPI; 1993-067439/08.  
XX  
PT New sulphated glyco-conjugate binding peptide(s) - from type 1 re  
peats of  
PT human thrombo:spondin, preventing interaction of the glyco-conjug  
ates  
PT with adhesion molecules, growth factors, etc.  
XX  
PS Disclosure; Page 13; 64pp; English.  
XX  
CC This peptide was obtained from repeat 2 of the adhesive glycoprot  
ein  
CC thrombospondin. It is one of five preferred peptides of the inven  
tion  
CC which all contain a subsequence WSXW (X=P, E, H, A, S) with a sub  
stantial  
CC lack of an electrical charge. Peptide #19 also includes the basic  
heparin  
CC -binding motif which enhanced activity approx. 3-fold for inhibit  
ing  
CC thrombospondin or laminin-binding to heparin but markedly decreas  
ed  
CC activity for inhibiting binding of both proteins to sulphatide. T  
he  
CC peptide acts as an inhibitor of heparin- or related sulphated  
CC glycoconjugate-binding to adhesion molecules, growth factors, etc  
. (Note:  
CC Revised entry submitted to correct the patent number format of US  
CC Government-owned NTIS applications to prevent clashes with ongoin  
g US  
CC granted patent numbers. For further information please visit the  
Derwent  
CC web site at [www.derwent.com/dwpi/updates/ntis\\_us.html](http://www.derwent.com/dwpi/updates/ntis_us.html).)  
XX  
SQ Sequence 17 AA;

us-10-030-735-1.rag

Query Match 100.0%; Score 108; DB 2; Length 17;  
Best Local Similarity 100.0%; Pred. No. 5.3e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; G  
aps 0;

Qy 1 KRFKQDGGWSHWSPWSS 17  
| | | | | | | | | | | | | | | | | |  
Db 1 KRFKQDGGWSHWSPWSS 17

RESULT 2

AAR55067

ID AAR55067 standard; peptide; 17 AA.

XX

AC AAR55067;

XX

DT 25-MAR-2003 (revised)

DT 16-DEC-1994 (first entry)

XX

DE Fibronectin gelatin binding domain inhibitory peptide.

XX

KW fibronectin; collagen-binding proteins; inhibit cell adhesion;

KW regulate cell matrix interactions; tumourigenesis; metastasis;

KW wound repair; homostasis; thrombospondin.

XX

OS Synthetic.

XX

PN WO9411395-A1.

XX

PD 26-MAY-1994.

XX

PF 09-NOV-1993; 93WO-US011104.

XX

PR 10-NOV-1992; 92US-00973235.

XX

PA (USSH ) US SEC DEPT HEALTH.

XX

PI Roberts DD, Krutzch HC, Sipes JM, Guo N, Negre E;

XX

DR WPI; 1994-183422/22.

XX

PT Peptides which bind to fibronectin and collagen-binding proteins  
- are

PT used to inhibit fibronectin dependent cell adhesion to collagen m  
atrices.

XX

PS Claim 1; Page 33; 51pp; English.



XX  
 CC Thrombospondin is a multi-functional protein capable of interacting with  
 CC numerous molecules, eg. fibronectin. Peptides have been designed  
 CC that are  
 CC derived from the second type I repeat of human endothelial cell  
 CC thrombospondin. The peptides can be used to bind to fibronectin or other  
 CC related collagen-binding proteins to inhibit fibronectin-dependent cell  
 CC adhesion to collagen matrices. The peptides interact directly with the  
 CC gelatin-binding domain of fibronectin and inhibits the fibronectin  
 CC function. (Updated on 25-MAR-2003 to correct PN field.) (Updated  
 CC on 25-MAR-2003 to correct PR field.)  
 XX  
 SQ Sequence 17 AA;

Query Match 100.0%; Score 108; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRFKQDGGWSHWSPWSS 17  
 |||||  
 Db 1 KRFKQDGGWSHWSPWSS 17

## RESULT 3

AAR69776

ID AAR69776 standard; peptide; 17 AA.

XX

AC AAR69776;

XX

DT 25-MAR-2003 (revised)

DT 13-OCT-1995 (first entry)

XX

DE Thrombospondin peptide #11 for activating latent TGF-beta.

XX

KW Thrombospondin type 1 repeat sequence; transforming growth factor -beta;

KW wound healing; fibrosis; endothelial cell proliferation.

XX

OS Synthetic.

XX

PN WO9505191-A1.

XX  
 PD 23-FEB-1995.  
 XX  
 PF 12-AUG-1994; 94WO-US009193.  
 XX  
 PR 13-AUG-1993; 93US-00106120.  
 PR 04-MAY-1994; 94US-00238169.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Murphy-Ullrich JE, Roberts DD, Kruttsch HC, Schultzcherry S;  
 XX  
 DR WPI; 1995-098579/13.  
 XX  
 PT Stimulating or inhibiting transforming growth factor-beta by cont  
 acting  
 PT with thrombo-spondin or an activating enzyme - used to enhance wo  
 und  
 PT healing or prevent fibrosis.  
 XX  
 PS Claim 4; Page 62; 67pp; English.  
 XX  
 CC The peptides AAR69766-79 are active peptide fragments based on th  
 e second  
 CC type 1 repeat sequence (amino acids 412-473) of thrombospondin (s  
 equence  
 CC not given in the specification) which can be used to convert late  
 nt  
 CC transforming growth factor-beta (TGF-b) to active TGF-b. thrombos  
 pondin,  
 CC or the activating peptides, can be used to stimulate TGF-b for us  
 e in  
 CC wound healing. Peptides (see AAR69780-90) which inhibit the stimu  
 lation  
 CC of TGF-b can be used to prevent fibrosis or block TGF-b mediated  
 CC endothelial cell proliferation. (Updated on 25-MAR-2003 to correc  
 t PN  
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)  
 XX  
 SQ Sequence 17 AA;

Query Match 100.0%; Score 108; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-08;  
 Matches 17; Conservative 0; Mismatches 0; Indels 0; G  
 aps 0;

Qy 1 KRFKQDGGWSHWSPWSS 17  
 |||||

Db

1 KRFBKQDGGWSHWSPWSS 17

## RESULT 4

AAB02656

ID AAB02656 standard; peptide; 17 AA.

XX

AC AAB02656;

XX

DT 18-AUG-2000 (first entry)

XX

DE Human thrombospondin derived peptide SEQ ID NO:19.

XX

KW Human; thrombospondin; inhibitor; type I repeat unit; identification;

KW extracellular matrix protein; heparin; proliferation; adhesion; motility;

KW extravasation; neovascularisation; metastatic tumour; breast carcinoma;

KW melanoma; cytostatic; heparin sulphate; FGF-2;

KW fibroblast growth factor 2; sulphate glycoconjugate; Kaposi's sarcoma;

KW haemangioma; diabetic retinopathy.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN US6051549-A.

XX

PD 18-APR-2000.

XX

PF 11-MAR-1998; 98US-00041119.

XX

PR 06-DEC-1991; 91US-00801812.

PR 21-MAR-1994; 94US-00215085.

PR 07-JUN-1995; 95US-00487568.

XX

PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

XX

PI Guo N, Kruttsch HC, Roberts DD;

XX

DR WPI; 2000-338337/29.

XX

PT New polypeptides binding to heparin or related sulfated glycoconjugates

PT are useful for inhibiting endothelial cell proliferation and tumor

PT growth.

XX

PS Example 1; Col 18; 133pp; English.

XX

CC The present invention describes a polypeptide (I) binding to heparin or

CC related sulphated glycoconjugates with high affinity. AAB02638 to

CC AAB02747 represent peptides derived from human thrombospondin, which are

CC used in the exemplification of the present invention. (I) can be used for

CC inhibiting heparin or heparin sulphate interaction in a sample, without

CC activating latent transforming growth factor (TGF)-beta, for inhibiting

CC interaction of heparin or heparin sulphate with FGF-2 in a sample, and

CC for inhibiting endothelial cell proliferation, and tumour growth, in

CC patients. The peptides bind heparin or related sulphate glycoconjugates

CC with high affinity. The peptides or their conjugates can be used in

CC blocking or modifying the action on cellular processes of heparin (e.g.

CC proliferation, adhesion, motility, extravasation and neovascularisation),

CC sulphatides, related sulphated glycoconjugates, fibronectin, and basic

CC fibroblast growth factor, involving malignant cell lines and normal

CC endothelial cells. The peptides can also be used in the treatment of

CC metastatic tumours, breast carcinomas, melanomas, Kaposi's sarcomas,

CC haemangiomas, diabetic retinopathies, and various pathological conditions

CC dependent on neovascularisation

XX

SQ Sequence 17 AA;

Query Match 100.0%; Score 108; DB 3; Length 17;

Best Local Similarity 100.0%; Pred. No. 5.3e-08;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KRFEKQDGGWSHWSPWSS 17

|||||||

Db

1 KRFKQDGGWSHWSPWSS 17

## RESULT 5

AAB35336

ID AAB35336 standard; peptide; 17 AA.

XX

AC AAB35336;

XX

DT 08-MAY-2001 (first entry)

XX

DE Alpha3beta1 integrin binding peptide #1.

XX

KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;

KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;

KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX

OS Synthetic.

XX

PN WO200105812-A2.

XX

PD 25-JAN-2001.

XX

PF 12-JUL-2000; 2000WO-US018986.

XX

PR 15-JUL-1999; 99US-0144549P.

XX

PA (USSH ) US DEPT HEALTH &amp; HUMAN SERVICES.

XX

PI Roberts DD, Kruttsch HC;

XX

DR WPI; 2001-182656/18.

XX

PT New peptides that bind to or are recognized by alpha3-beta1 integrins,

PT useful for inhibiting cell adhesion to extracellular matrix, cell

PT motility and proliferation and for treating rheumatoid arthritis and

PT cancer.

XX

PS Disclosure; Page 3; 84pp; English.

XX

CC The present invention provides a number of peptides which bind to

CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion

us-10-030-735-1.rag

CC and motility, and in the treatment of cancer, diabetic retinopath  
y,  
CC rheumatoid arthritis, macular degeneration, atherosclerosis, psor  
iasis  
CC and restenosis. The present sequence is an example of one of the  
peptides  
CC of the invention  
XX  
SQ Sequence 17 AA;

Query Match 100.0%; Score 108; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 5.3e-08;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; G  
aps 0;

Qy 1 KRFKQDGGWSHWSPWSS 17  
| | | | | | | | | | | | | | | | | |  
Db 1 KRFKQDGGWSHWSPWSS 17

# RESULT 6

ABU13826

ID ABU13826 standard; peptide; 17 AA.

XX

AC ABU13826;

XX

DT 25-FEB-2003 (first entry)

XX

DE Thrombospondin dependent TGF-beta activation inhibitory peptide #  
15.

XX

KW Thrombospondin dependent TGF-beta activation; TGF-beta; TSP1;

KW thrombospondin; transforming growth factor beta; litholytic;

KW nephrotropic; TGF-beta inhibitor; kidney disease; renal system di  
sorder;

KW mesangial proliferative glomerulonephritis; scarring glomerular d  
isease;

KW microaneurysm formation.

XX

OS Homo sapiens.

XX

PN US6458767-B1.

XX

PD 01-OCT-2002.

XX

PF 28-MAY-1999; 99US-00321932.

XX

PR 04-MAY-1994; 94US-00238169.

SEQ ID NO: 18

10/6301735

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:39:01 ; Search time 33.8326 Seconds  
(without alignments)  
93.153 Million cell updates/sec

Title: US-10-030-735-18  
Perfect score: 61  
Sequence: 1 FQGV LQNLR FVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Match	Query	Length	DB	ID	Description
--------	-------	-------	-------	--------	----	----	-------------

1	58	95.1	240	12	US-10-419-462-40	Sequence 40, Appl
2	58	95.1	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	58	95.1	831	12	US-09-939-853A-97	Sequence 97, Appl
4	58	95.1	831	12	US-09-939-853A-98	Sequence 98, Appl
5	58	95.1	1152	9	US-09-919-603-1	Sequence 1, Appli
6	58	95.1	1170	12	US-10-211-462-38	Sequence 38, Appl
7	58	95.1	1170	12	US-10-231-956A-482	Sequence 482, App
8	58	95.1	1170	12	US-10-419-462-38	Sequence 38, Appl
9	58	95.1	1170	14	US-10-020-141-12	Sequence 12, Appl
10	58	95.1	1170	14	US-10-017-721-2	Sequence 2, Appli
11	58	95.1	1170	14	US-10-021-660-114	Sequence 114, App
12	58	95.1	1170	14	US-10-008-093-2	Sequence 2, Appli
13	58	95.1	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	42	68.9	15	15	US-10-285-394-153	Sequence 153, App
15	40	65.6	16	9	US-09-822-682-6	Sequence 6, Appli
16	40	65.6	151	9	US-09-925-299-1251	Sequence 1251, Ap
17	40	65.6	151	10	US-09-925-299-1251	Sequence 1251, Ap
18	40	65.6	710	12	US-10-282-122A-75098	Sequence 75098, A
19	40	65.6	1168	9	US-09-919-603-2	Sequence 2, Appli
20	40	65.6	1172	9	US-09-919-770-4	Sequence 4, Appli
21	40	65.6	1172	9	US-09-822-682-2	Sequence 2, Appli
22	40	65.6	1172	10	US-09-866-050A-658	Sequence 658, App
23	40	65.6	1172	14	US-10-060-036-171	Sequence 171, App
24	40	65.6	1172	14	US-10-020-141-14	Sequence 14, Appl
25	40	65.6	1172	14	US-10-017-724-2	Sequence 2, Appli
26	40	65.6	1172	14	US-10-301-822-203	Sequence 203, App
27	40	65.6	1172	14	US-10-008-093-4	Sequence 4, Appli
28	40	65.6	1172	14	US-10-007-781-2	Sequence 2, Appli
29	40	65.6	1172	15	US-10-372-683-16	Sequence 16, Appl
30	38	62.3	539	15	US-10-422-536-16	Sequence 16, Appl
31	38	62.3	759	12	US-10-282-122A-77853	Sequence 77853, A
32	38	62.3	1300	10	US-09-896-923-3	Sequence 3, Appli
33	38	62.3	1300	15	US-10-223-650-3	Sequence 3, Appli
34	38	62.3	1751	15	US-10-435-766-103	Sequence 103, App
35	37	60.7	87	12	US-10-424-599-254511	Sequence 254511,
36	37	60.7	325	12	US-10-425-114-71821	Sequence 71821, A
37	37	60.7	466	12	US-10-424-599-232021	Sequence 232021,
38	36	59.0	215	14	US-10-029-386-33750	Sequence 33750, A
39	36	59.0	226	9	US-09-925-300-1217	Sequence 1217, Ap
40	36	59.0	390	14	US-10-420-845-6	Sequence 6, Appli
41	36	59.0	559	14	US-10-284-499-47	Sequence 47, Appl
42	36	59.0	651	14	US-10-420-845-5	Sequence 5, Appli
43	36	59.0	856	12	US-10-282-122A-65000	Sequence 65000, A
44	36	59.0	1070	14	US-10-420-845-22	Sequence 22, Appl
45	36	59.0	1457	9	US-09-772-316-2	Sequence 2, Appli
46	36	59.0	1457	15	US-10-600-058-2	Sequence 2, Appli
47	36	59.0	1504	9	US-09-932-145-7	Sequence 7, Appli
48	36	59.0	1504	14	US-10-284-499-1	Sequence 1, Appli
49	36	59.0	1504	15	US-10-395-812-7	Sequence 7, Appli
50	35	57.4	72	12	US-10-424-599-147545	Sequence 147545,
51	35	57.4	106	12	US-10-424-599-242530	Sequence 242530,
52	35	57.4	226	9	US-09-895-913A-362	Sequence 362, App
53	35	57.4	247	9	US-09-815-242-11342	Sequence 11342, A
54	35	57.4	247	9	US-09-815-242-11514	Sequence 11514, A
55	35	57.4	247	12	US-10-282-122A-58737	Sequence 58737, A
56	35	57.4	293	15	US-10-369-493-10330	Sequence 10330, A



57	35	57.4	410	15	US-10-369-493-8711	Sequence 8711, Ap
58	35	57.4	459	12	US-10-282-122A-67080	Sequence 67080, A
59	35	57.4	471	12	US-10-282-122A-49774	Sequence 49774, A
60	35	57.4	481	15	US-10-369-493-16410	Sequence 16410, A
61	35	57.4	922	14	US-10-128-714-3524	Sequence 3524, Ap
62	35	57.4	949	14	US-10-128-714-8524	Sequence 8524, Ap
63	35	57.4	1829	15	US-10-435-766-13	Sequence 13, Appl
64	35	57.4	3500	14	US-10-153-219-2	Sequence 2, Appli
65	35	57.4	3537	14	US-10-153-219-15	Sequence 15, Appl
66	34	55.7	74	12	US-10-424-599-187467	Sequence 187467,
67	34	55.7	84	12	US-10-424-599-160481	Sequence 160481,
68	34	55.7	86	12	US-10-424-599-164070	Sequence 164070,
69	34	55.7	92	12	US-10-424-599-241959	Sequence 241959,
70	34	55.7	104	10	US-09-764-891-4845	Sequence 4845, Ap
71	34	55.7	104	14	US-10-091-572-391	Sequence 391, App
72	34	55.7	213	12	US-10-282-122A-48068	Sequence 48068, A
73	34	55.7	213	12	US-10-282-122A-50493	Sequence 50493, A
74	34	55.7	282	12	US-10-424-599-244115	Sequence 244115,
75	34	55.7	314	15	US-10-369-493-12457	Sequence 12457, A
76	34	55.7	374	12	US-10-424-599-181131	Sequence 181131,
77	34	55.7	465	10	US-09-769-787-164	Sequence 164, App
78	34	55.7	482	12	US-10-425-114-42018	Sequence 42018, A
79	34	55.7	565	12	US-10-282-122A-50809	Sequence 50809, A
80	34	55.7	634	12	US-10-282-122A-46137	Sequence 46137, A
81	34	55.7	826	12	US-10-424-599-166203	Sequence 166203,
82	34	55.7	949	14	US-10-148-351-60	Sequence 60, Appl
83	33	54.1	45	12	US-10-424-599-275218	Sequence 275218,
84	33	54.1	50	9	US-09-971-980-64	Sequence 64, Appl
85	33	54.1	51	12	US-10-424-599-198354	Sequence 198354,
86	33	54.1	59	11	US-09-864-408A-8466	Sequence 8466, Ap
87	33	54.1	64	9	US-09-971-980-62	Sequence 62, Appl
88	33	54.1	75	12	US-10-424-599-271195	Sequence 271195,
89	33	54.1	75	12	US-10-424-599-283340	Sequence 283340,
90	33	54.1	83	12	US-10-424-599-150515	Sequence 150515,
91	33	54.1	89	12	US-10-424-599-153242	Sequence 153242,
92	33	54.1	125	12	US-10-424-599-259520	Sequence 259520,
93	33	54.1	131	12	US-10-424-599-272021	Sequence 272021,
94	33	54.1	163	12	US-10-424-599-237046	Sequence 237046,
95	33	54.1	178	15	US-10-108-260A-3713	Sequence 3713, Ap
96	33	54.1	188	12	US-10-424-599-147763	Sequence 147763,
97	33	54.1	216	15	US-10-264-049-4189	Sequence 4189, Ap
98	33	54.1	365	16	US-10-344-738-76	Sequence 76, Appl
99	33	54.1	372	15	US-10-369-493-6644	Sequence 6644, Ap
100	33	54.1	387	9	US-09-738-626-3890	Sequence 3890, Ap
101	33	54.1	426	12	US-10-282-122A-52492	Sequence 52492, A
102	33	54.1	426	12	US-10-282-122A-73519	Sequence 73519, A
103	33	54.1	429	14	US-10-152-254-6	Sequence 6, Appli
104	33	54.1	445	12	US-10-282-122A-57442	Sequence 57442, A
105	33	54.1	453	9	US-09-815-242-10204	Sequence 10204, A
106	33	54.1	453	12	US-10-282-122A-42861	Sequence 42861, A
107	33	54.1	453	12	US-10-282-122A-60049	Sequence 60049, A
108	33	54.1	453	12	US-10-282-122A-75319	Sequence 75319, A
109	33	54.1	453	12	US-10-282-122A-76238	Sequence 76238, A
110	33	54.1	454	9	US-09-815-242-13928	Sequence 13928, A
111	33	54.1	464	12	US-10-282-122A-78429	Sequence 78429, A
112	33	54.1	586	12	US-10-282-122A-45424	Sequence 45424, A
113	33	54.1	586	12	US-10-282-122A-46433	Sequence 46433, A

114	33	54.1	597	12	US-10-282-122A-76883	Sequence 76883, A
115	33	54.1	764	9	US-09-895-913A-328	Sequence 328, App
116	33	54.1	764	12	US-10-282-122A-58705	Sequence 58705, A
117	33	54.1	1360	15	US-10-334-143-17	Sequence 17, Appl
118	32	52.5	48	12	US-10-424-599-197249	Sequence 197249,
119	32	52.5	55	13	US-10-016-634A-119	Sequence 119, App
120	32	52.5	69	12	US-10-424-599-206194	Sequence 206194,
121	32	52.5	78	12	US-10-424-599-195073	Sequence 195073,
122	32	52.5	81	14	US-10-029-386-29347	Sequence 29347, A
123	32	52.5	89	12	US-10-424-599-165087	Sequence 165087,
124	32	52.5	108	15	US-10-369-493-13199	Sequence 13199, A
125	32	52.5	113	12	US-10-424-599-211875	Sequence 211875,
126	32	52.5	113	14	US-10-055-882-4	Sequence 4, Appli
127	32	52.5	119	11	US-09-833-245-75	Sequence 75, Appl
128	32	52.5	130	9	US-09-949-192-41	Sequence 41, Appl
129	32	52.5	159	10	US-09-809-391-499	Sequence 499, App
130	32	52.5	159	10	US-09-882-171-499	Sequence 499, App
131	32	52.5	177	12	US-10-424-599-145722	Sequence 145722,
132	32	52.5	186	12	US-10-425-114-71673	Sequence 71673, A
133	32	52.5	191	12	US-10-072-012-26	Sequence 26, Appl
134	32	52.5	215	12	US-10-072-012-28	Sequence 28, Appl
135	32	52.5	226	12	US-10-425-114-41620	Sequence 41620, A
136	32	52.5	228	12	US-10-282-122A-49254	Sequence 49254, A
137	32	52.5	235	12	US-10-424-599-181083	Sequence 181083,
138	32	52.5	242	12	US-10-424-599-228787	Sequence 228787,
139	32	52.5	245	12	US-10-425-114-63529	Sequence 63529, A
140	32	52.5	252	12	US-10-425-114-66807	Sequence 66807, A
141	32	52.5	261	12	US-10-425-114-66215	Sequence 66215, A
142	32	52.5	266	12	US-10-425-114-72134	Sequence 72134, A
143	32	52.5	268	12	US-10-425-114-63145	Sequence 63145, A
144	32	52.5	269	12	US-10-425-114-50585	Sequence 50585, A
145	32	52.5	269	12	US-10-425-114-72162	Sequence 72162, A
146	32	52.5	269	12	US-10-425-114-72180	Sequence 72180, A
147	32	52.5	270	12	US-10-425-114-60079	Sequence 60079, A
148	32	52.5	274	12	US-10-425-114-66637	Sequence 66637, A
149	32	52.5	275	12	US-10-425-114-66540	Sequence 66540, A
150	32	52.5	275	12	US-10-425-114-66844	Sequence 66844, A

#### ALIGNMENTS

#### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 95.1%; Score 58; DB 12; Length 240;  
Best Local Similarity 91.7%; Pred. No. 0.011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGV LQN LRFVF 12  
||| ||| : |||  
Db 190 FQGV LQN VRFVF 201

RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 95.1%; Score 58; DB 9; Length 466;  
Best Local Similarity 91.7%; Pred. No. 0.022;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGV LQN LRFVF 12  
||| ||| : |||  
Db 261 FQGV LQN VRFVF 272

RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1e1 Proteins and Nucleic Acids Encoding  
Same

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:04:45 ; Search time 12.9515 Seconds  
(without alignments)  
47.833 Million cell updates/sec

Title: US-10-030-735-18  
Perfect score: 61  
Sequence: 1 FQGV LQNLRFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	58	95.1	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	40	65.6	1172	1	US-08-313-288B-19	Sequence 19, Appl
3	38	62.3	539	4	US-09-800-170-16	Sequence 16, Appl
4	38	62.3	1019	1	US-08-271-364A-7	Sequence 7, Appli
5	38	62.3	1019	2	US-08-222-715B-26	Sequence 26, Appl
6	37	60.7	432	4	US-09-489-039A-7909	Sequence 7909, Ap
7	36	59.0	390	3	US-08-650-766-7	Sequence 7, Appli
8	36	59.0	390	3	US-08-922-635-6	Sequence 6, Appli
9	36	59.0	390	4	US-09-389-487-7	Sequence 7, Appli
10	36	59.0	559	4	US-09-364-206-47	Sequence 47, Appl
11	36	59.0	651	3	US-08-650-766-6	Sequence 6, Appli

12	36	59.0	651	3	US-08-922-635-5	Sequence 5, Appli
13	36	59.0	651	4	US-09-389-487-6	Sequence 6, Appli
14	36	59.0	1022	1	US-08-271-364A-8	Sequence 8, Appli
15	36	59.0	1022	2	US-08-222-715B-27	Sequence 27, Appl
16	36	59.0	1070	3	US-08-922-635-22	Sequence 22, Appl
17	36	59.0	1504	4	US-09-364-206-2	Sequence 2, Appli
18	35	57.4	903	4	US-09-252-991A-28775	Sequence 28775, A
19	35	57.4	1350	2	US-08-319-866-9	Sequence 9, Appli
20	35	57.4	2475	3	US-09-413-814-48	Sequence 48, Appl
21	34	55.7	415	4	US-09-489-039A-10457	Sequence 10457, A
22	34	55.7	740	4	US-09-134-000C-6441	Sequence 6441, Ap
23	34	55.7	1288	3	US-08-762-428A-6	Sequence 6, Appli
24	33	54.1	169	4	US-09-134-000C-3511	Sequence 3511, Ap
25	33	54.1	445	4	US-09-107-532A-5506	Sequence 5506, Ap
26	33	54.1	476	4	US-09-489-039A-8120	Sequence 8120, Ap
27	33	54.1	1333	4	US-09-347-878-20	Sequence 20, Appl
28	32	52.5	159	4	US-09-149-476-499	Sequence 499, App
29	32	52.5	175	4	US-09-252-991A-18826	Sequence 18826, A
30	32	52.5	238	6	5405943-2	Patent No. 5405943
31	32	52.5	275	4	US-09-134-001C-3732	Sequence 3732, Ap
32	32	52.5	357	4	US-09-556-601-26	Sequence 26, Appl
33	32	52.5	390	4	US-09-328-352-4891	Sequence 4891, Ap
34	32	52.5	436	6	5405943-4	Patent No. 5405943
35	32	52.5	440	1	US-08-307-499-15	Sequence 15, Appl
36	32	52.5	440	3	US-09-299-268-15	Sequence 15, Appl
37	32	52.5	457	4	US-09-198-452A-553	Sequence 553, App
38	32	52.5	538	4	US-09-134-000C-4331	Sequence 4331, Ap
39	32	52.5	616	4	US-09-543-681A-4421	Sequence 4421, Ap
40	32	52.5	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
41	32	52.5	1306	3	US-08-999-774A-13	Sequence 13, Appl
42	32	52.5	1307	1	US-08-395-246C-2	Sequence 2, Appli
43	31	50.8	24	1	US-08-019-073-15	Sequence 15, Appl
44	31	50.8	24	5	PCT-US94-01768-15	Sequence 15, Appl
45	31	50.8	85	2	US-08-480-229C-5	Sequence 5, Appli
46	31	50.8	85	2	US-08-659-235C-5	Sequence 5, Appli
47	31	50.8	102	4	US-09-732-210-1730	Sequence 1730, Ap
48	31	50.8	156	2	US-08-162-402B-18	Sequence 18, Appl
49	31	50.8	175	4	US-09-479-313B-4	Sequence 4, Appli
50	31	50.8	178	4	US-09-252-991A-22198	Sequence 22198, A
51	31	50.8	206	4	US-09-252-991A-20515	Sequence 20515, A
52	31	50.8	226	3	US-09-176-657-3	Sequence 3, Appli
53	31	50.8	226	4	US-09-421-299-3	Sequence 3, Appli
54	31	50.8	295	3	US-09-150-133-13	Sequence 13, Appl
55	31	50.8	295	3	US-09-150-141-13	Sequence 13, Appl
56	31	50.8	295	3	US-09-374-493-13	Sequence 13, Appl
57	31	50.8	295	3	US-09-374-824-13	Sequence 13, Appl
58	31	50.8	295	3	US-09-374-492-13	Sequence 13, Appl
59	31	50.8	295	4	US-09-785-343-13	Sequence 13, Appl
60	31	50.8	297	4	US-09-489-039A-10616	Sequence 10616, A
61	31	50.8	341	4	US-09-489-039A-12069	Sequence 12069, A
62	31	50.8	342	4	US-09-134-001C-4071	Sequence 4071, Ap
63	31	50.8	342	4	US-09-252-991A-16724	Sequence 16724, A
64	31	50.8	345	4	US-09-252-991A-28350	Sequence 28350, A
65	31	50.8	409	4	US-09-489-039A-14217	Sequence 14217, A
66	31	50.8	420	4	US-09-252-991A-19389	Sequence 19389, A
67	31	50.8	448	4	US-09-107-532A-6632	Sequence 6632, Ap
68	31	50.8	449	4	US-09-489-039A-8009	Sequence 8009, Ap

69	31	50.8	468	1	US-08-204-656B-2	Sequence 2, Appli
70	31	50.8	468	1	US-08-204-656B-4	Sequence 4, Appli
71	31	50.8	468	1	US-08-204-656B-6	Sequence 6, Appli
72	31	50.8	468	1	US-08-204-656B-8	Sequence 8, Appli
73	31	50.8	468	1	US-08-470-702-6	Sequence 6, Appli
74	31	50.8	468	1	US-08-470-702-7	Sequence 7, Appli
75	31	50.8	468	1	US-08-470-702-8	Sequence 8, Appli
76	31	50.8	468	1	US-08-470-702-9	Sequence 9, Appli
77	31	50.8	468	1	US-08-467-831-6	Sequence 6, Appli
78	31	50.8	468	1	US-08-467-831-7	Sequence 7, Appli
79	31	50.8	468	1	US-08-467-831-8	Sequence 8, Appli
80	31	50.8	468	1	US-08-467-831-9	Sequence 9, Appli
81	31	50.8	468	4	US-09-543-681A-8164	Sequence 8164, Ap
82	31	50.8	470	4	US-08-687-691B-4	Sequence 4, Appli
83	31	50.8	479	4	US-09-328-352-4361	Sequence 4361, Ap
84	31	50.8	504	1	US-08-135-511-30	Sequence 30, Appl
85	31	50.8	504	1	US-08-187-453-30	Sequence 30, Appl
86	31	50.8	512	4	US-09-252-991A-25181	Sequence 25181, A
87	31	50.8	715	2	US-08-484-993B-10	Sequence 10, Appl
88	31	50.8	715	2	US-08-484-158B-10	Sequence 10, Appl
89	31	50.8	715	2	US-08-484-596A-10	Sequence 10, Appl
90	31	50.8	715	2	US-08-480-150A-10	Sequence 10, Appl
91	31	50.8	715	3	US-08-458-731-10	Sequence 10, Appl
92	31	50.8	715	3	US-08-149-223A-10	Sequence 10, Appl
93	31	50.8	810	2	US-08-820-170A-34	Sequence 34, Appl
94	31	50.8	810	3	US-09-055-699-34	Sequence 34, Appl
95	31	50.8	810	3	US-09-273-565-34	Sequence 34, Appl
96	31	50.8	810	4	US-09-565-538-34	Sequence 34, Appl
97	31	50.8	810	4	US-09-661-468-34	Sequence 34, Appl
98	31	50.8	810	4	US-09-976-165-34	Sequence 34, Appl
99	31	50.8	894	4	US-09-134-000C-5731	Sequence 5731, Ap
100	31	50.8	900	4	US-09-328-352-8072	Sequence 8072, Ap
101	31	50.8	902	4	US-09-107-532A-3837	Sequence 3837, Ap
102	31	50.8	944	4	US-09-134-000C-5578	Sequence 5578, Ap
103	31	50.8	1996	2	US-08-804-227C-9	Sequence 9, Appli
104	31	50.8	1996	2	US-08-804-198-3	Sequence 3, Appli
105	31	50.8	4545	2	US-08-804-227C-14	Sequence 14, Appl
106	31	50.8	4550	2	US-08-804-227C-8	Sequence 8, Appli
107	31	50.8	4550	2	US-08-804-198-2	Sequence 2, Appli
108	30	49.2	35	3	US-08-942-686-5	Sequence 5, Appli
109	30	49.2	40	1	US-07-868-353A-3	Sequence 3, Appli
110	30	49.2	40	2	US-08-407-804-3	Sequence 3, Appli
111	30	49.2	40	3	US-09-124-807-3	Sequence 3, Appli
112	30	49.2	51	4	US-09-461-325-184	Sequence 184, App
113	30	49.2	51	4	US-10-012-542-184	Sequence 184, App
114	30	49.2	52	3	US-08-917-299-25	Sequence 25, Appl
115	30	49.2	52	3	US-09-422-662-25	Sequence 25, Appl
116	30	49.2	52	4	US-09-730-763-25	Sequence 25, Appl
117	30	49.2	52	4	US-09-429-370-25	Sequence 25, Appl
118	30	49.2	56	4	US-09-205-258-984	Sequence 984, App
119	30	49.2	72	4	US-09-461-325-407	Sequence 407, App
120	30	49.2	72	4	US-10-012-542-407	Sequence 407, App
121	30	49.2	104	4	US-09-230-485-9	Sequence 9, Appli
122	30	49.2	127	4	US-09-614-912-32	Sequence 32, Appl
123	30	49.2	163	4	US-09-489-039A-10183	Sequence 10183, A
124	30	49.2	175	3	US-09-230-637-24	Sequence 24, Appl
125	30	49.2	191	4	US-09-205-258-981	Sequence 981, App

126	30	49.2	197	4	US-09-252-991A-23101	Sequence 23101, A
127	30	49.2	208	4	US-09-134-001C-3478	Sequence 3478, Ap
128	30	49.2	241	4	US-09-252-991A-25199	Sequence 25199, A
129	30	49.2	258	4	US-09-540-236-2125	Sequence 2125, Ap
130	30	49.2	306	4	US-09-634-238-367	Sequence 367, Appl
131	30	49.2	309	4	US-09-252-991A-27914	Sequence 27914, A
132	30	49.2	323	4	US-09-252-991A-21180	Sequence 21180, A
133	30	49.2	338	4	US-09-252-991A-32574	Sequence 32574, A
134	30	49.2	346	4	US-09-083-268-16	Sequence 16, Appl
135	30	49.2	350	1	US-07-868-353A-14	Sequence 14, Appl
136	30	49.2	350	2	US-08-407-804-23	Sequence 23, Appl
137	30	49.2	350	3	US-09-124-807-23	Sequence 23, Appl
138	30	49.2	350	4	US-09-252-991A-29676	Sequence 29676, A
139	30	49.2	354	1	US-07-868-353A-12	Sequence 12, Appl
140	30	49.2	354	1	US-07-868-353A-13	Sequence 13, Appl
141	30	49.2	354	1	US-07-868-353A-15	Sequence 15, Appl
142	30	49.2	354	2	US-08-407-804-21	Sequence 21, Appl
143	30	49.2	354	2	US-08-407-804-22	Sequence 22, Appl
144	30	49.2	354	2	US-08-407-804-24	Sequence 24, Appl
145	30	49.2	354	3	US-09-124-807-21	Sequence 21, Appl
146	30	49.2	354	3	US-09-124-807-22	Sequence 22, Appl
147	30	49.2	354	3	US-09-124-807-24	Sequence 24, Appl
148	30	49.2	358	4	US-09-252-991A-27541	Sequence 27541, A
149	30	49.2	389	4	US-09-464-035A-7	Sequence 7, Appli
150	30	49.2	415	4	US-09-489-039A-10601	Sequence 10601, A

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-313-288B-20

Query Match 95.1%; Score 58; DB 1; Length 1170;  
 Best Local Similarity 91.7%; Pred. No. 0.017;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQLQNLRFVF 12  
 |||||:||||  
 Db 208 FQGVQLQNVRFVF 219

RESULT 2

US-08-313-288B-19  
 ; Sequence 19, Application US/08313288B  
 ; Patent No. 5750502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jessell, Thomas M. and Avihu Klar  
 ; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
 ; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/313,288B  
 ; FILING DATE: January 5, 1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:



OM protein - protein search, using sw model

Run on: April 7, 2004, 18:03:50 ; Search time 9.67401 Seconds  
 (without alignments)  
 119.320 Million cell updates/sec

Title: US-10-030-735-18  
 Perfect score: 61  
 Sequence: 1 FQGVLQNLRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	58	95.1	229	2	S57957	thrombospondin 1 -	
2	58	95.1	1170	1	TSHUP1	thrombospondin 1 p	
3	58	95.1	1170	2	A40558	thrombospondin 1 p	
4	41	67.2	145	2	T20985	hypothetical prote	
5	41	67.2	162	2	B88349	protein F15D4.3 [i	
6	41	67.2	467	2	D84938	H+-transporting tw	
7	40	65.6	1172	1	TSHUP2	thrombospondin 2 p	
8	40	65.6	1172	2	A42587	thrombospondin 2 p	
9	40	65.6	1178	1	A39804	thrombospondin pre	
10	38	62.3	212	2	D81929	probable imidazole	
11	38	62.3	759	2	AC0368	probable autotrans	
12	38	62.3	1312	2	S68593	DNA-directed DNA p	
13	38	62.3	1670	2	S71551	DNA-directed DNA p	

14	37	60.7	337	2	E97882	hypothetical prote
15	37	60.7	471	2	S45068	53K glycoprotein -
16	36	59.0	199	1	G64070	imidazoleglycerol-
17	36	59.0	385	2	S56224	hypothetical prote
18	36	59.0	492	2	S32491	testosterone 7alph
19	36	59.0	511	2	S77350	hypothetical prote
20	36	59.0	750	1	HYHUN	neprilysin (EC 3.4
21	36	59.0	750	1	HYRTN	neprilysin (EC 3.4
22	36	59.0	751	1	HYRBN	neprilysin (EC 3.4
23	36	59.0	876	2	B96693	probable receptor
24	36	59.0	993	2	T17230	hypothetical prote
25	36	59.0	1038	2	T02634	rep protein homolo
26	36	59.0	1702	2	S42459	DNA-directed DNA p
27	35	57.4	247	1	A64590	probable 3-oxoacyl
28	35	57.4	247	2	B71923	3-oxoacyl-[acyl-ca
29	35	57.4	392	2	G69896	conserved hypothet
30	35	57.4	431	2	T32359	hypothetical prote
31	35	57.4	455	2	D83264	hypothetical prote
32	35	57.4	465	2	AC0347	probable membrane
33	35	57.4	531	2	JC5172	probable methylgal
34	35	57.4	834	2	T39891	probable integral
35	35	57.4	893	2	T36795	probable penicilli
36	35	57.4	1054	2	H69377	reverse gyrase (to
37	35	57.4	1350	2	T13254	nitric-oxide synth
38	34.5	56.6	79	2	C69036	hypothetical prote
39	34	55.7	133	2	B89969	enterotoxin Yent1
40	34	55.7	265	2	F69742	hypothetical prote
41	34	55.7	290	2	AD1858	indole-3-glycerol
42	34	55.7	304	2	T34271	hypothetical prote
43	34	55.7	354	2	T20169	hypothetical prote
44	34	55.7	457	2	JC4541	NADH oxidase (H2O-
45	34	55.7	498	2	B89832	hypothetical prote
46	34	55.7	562	2	AD0328	probable potassium
47	34	55.7	585	2	S43572	C05B5.5 protein (c
48	34	55.7	585	2	E88571	protein C05B5.5 [i
49	34	55.7	688	2	S55349	potassium channel
50	34	55.7	740	2	G95153	neuraminidase, pro
51	34	55.7	927	2	T43110	lacticin 481/lacto
52	34	55.7	1001	2	T28897	hypothetical prote
53	34	55.7	1288	2	T42756	5-oxoprolinase (AT
54	33.5	54.9	827	2	A95877	hypothetical prote
55	33	54.1	76	2	T29579	hypothetical prote
56	33	54.1	151	2	C57253	tRNA-pseudouridine
57	33	54.1	186	2	B75421	probable pilin, ty
58	33	54.1	193	2	E86716	acetyl transferase
59	33	54.1	194	2	T04180	ribosomal protein
60	33	54.1	206	2	S70004	hypothetical prote
61	33	54.1	222	2	A26489	placental lactogen
62	33	54.1	360	2	AB0104	LacI-family regula
63	33	54.1	372	2	T24392	hypothetical prote
64	33	54.1	399	2	AG2221	hypothetical prote
65	33	54.1	447	2	A86244	Barley Mlo protein
66	33	54.1	453	2	H64974	hypothetical prote
67	33	54.1	453	2	A98990	hypothetical prote
68	33	54.1	453	2	D85835	hypothetical prote
69	33	54.1	453	2	AE0774	probable protease
70	33	54.1	464	2	AG0347	probable proteinas

71	33	54.1	465	2	S76464	hypothetical prote
72	33	54.1	474	2	S07754	NADH2 dehydrogenas
73	33	54.1	492	2	A31887	testosterone 7alph
74	33	54.1	492	2	A34272	testosterone 7alph
75	33	54.1	555	2	T23531	hypothetical prote
76	33	54.1	563	2	T36704	probable dehydroge
77	33	54.1	597	2	B82881	hypothetical prote
78	33	54.1	672	2	A65024	Hydrogenase-4 comp
79	33	54.1	672	2	H91046	hydrogenase 4 memb
80	33	54.1	672	2	D85891	hydrogenase 4 memb
81	33	54.1	681	2	T15590	hypothetical prote
82	33	54.1	764	2	B64570	phenylalanine-tRNA
83	33	54.1	839	2	T21207	hypothetical prote
84	33	54.1	946	2	S71168	Ca2+-transporting
85	33	54.1	1020	2	D86402	protein envelope C
86	33	54.1	1020	2	T51925	Ca2+-transporting
87	33	54.1	1020	2	T51926	Ca2+-transporting
88	33	54.1	1069	2	C85349	Ca2+-transporting
89	33	54.1	1080	2	A35088	phycobilisome link
90	33	54.1	1093	2	T08551	Ca2+-transporting
91	33	54.1	1132	2	AD1809	phycobilisome core
92	33	54.1	1177	2	I64233	hypothetical prote
93	33	54.1	1207	2	C70013	conserved hypothet
94	33	54.1	1333	1	XOHUDH	xanthine dehydroge
95	33	54.1	1829	2	AE1864	hypothetical prote
96	33	54.1	1839	2	S77626	mannuronan C-5-epi
97	33	54.1	3587	2	I40486	surfactin syntheta
98	33	54.1	6420	2	T30283	polyketide synthas
99	32	52.5	54	2	S35697	leukocidin chain F
100	32	52.5	93	2	T31048	hypothetical prote
101	32	52.5	111	2	S52596	probable membrane
102	32	52.5	159	2	F95001	transcription regu
103	32	52.5	159	2	H95234	transcription regu
104	32	52.5	159	2	E97873	hypothetical prote
105	32	52.5	159	2	A98099	hypothetical prote
106	32	52.5	160	2	S65978	spore coat protein
107	32	52.5	172	2	G95289	hypothetical prote
108	32	52.5	263	2	AF0675	probable regulator
109	32	52.5	270	2	T06984	hypothetical prote
110	32	52.5	276	1	FLQL2C	flagellin, 28.5K -
111	32	52.5	286	2	C49238	gamma-hemolysin co
112	32	52.5	295	1	S76016	hypothetical prote
113	32	52.5	298	2	T12084	hypothetical prote
114	32	52.5	307	2	H88950	protein C38C3.4 [i
115	32	52.5	310	2	S68225	synergohymenotropi
116	32	52.5	311	2	F70184	ribose/galactose A
117	32	52.5	312	2	T00160	leukocidin chain S
118	32	52.5	312	2	S32211	leucocidin chain S
119	32	52.5	315	2	PC4078	hlgC-like protein
120	32	52.5	315	2	A49234	leucocidin R S com
121	32	52.5	315	2	JN0626	leukocidin chain S
122	32	52.5	315	2	E90043	gamma-hemolysin co
123	32	52.5	326	2	S28706	hypothetical prote
124	32	52.5	326	2	H83832	hypothetical prote
125	32	52.5	331	2	T21156	hypothetical prote
126	32	52.5	348	2	F70398	p-aminobenzoate sy
127	32	52.5	349	2	C96738	unknown protein F3

128	32	52.5	351	2	D89991	hypothetical prote
129	32	52.5	355	2	T05390	probable cysteine
130	32	52.5	355	2	T50479	G protein alpha ch
131	32	52.5	363	2	AE1712	PTS system, fructo
132	32	52.5	364	2	AG1341	PTS system, fructo
133	32	52.5	369	2	S74017	hypothetical prote
134	32	52.5	405	2	G82165	glucose-1-phosphat
135	32	52.5	406	2	G02022	tryptophan oxygena
136	32	52.5	445	2	E86382	hypothetical prote
137	32	52.5	448	2	S73718	hypothetical prote
138	32	52.5	454	2	D86793	drug-export protei
139	32	52.5	457	2	A54604	regulatory protein
140	32	52.5	469	2	T33595	hypothetical prote
141	32	52.5	514	2	B69205	cobyric acid synth
142	32	52.5	553	2	AH1830	serine/threonine k
143	32	52.5	571	2	H70623	probable potassium
144	32	52.5	587	2	T16867	probable cytochrom
145	32	52.5	589	2	T50698	probable mandeloni
146	32	52.5	620	2	F72069	conserved hypothet
147	32	52.5	620	2	D86554	CT425 hypothetical
148	32	52.5	644	2	A42220	helix-loop-helix p
149	32	52.5	650	2	S75072	probable phytoene
150	32	52.5	652	2	T03504	probable DNA topoi

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 95.1%; Score 58; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.00081;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FQGV LQN LRFVF 12
          ||| ||| ||| : |||
Db     190 FQGV LQN VRFVF 201

```

##### RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.

A;Reference number: A26155; MUID:87057617; PMID:2430973

A;Accession: A26155

A;Molecule type: mRNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.

A;Reference number: A34274; MUID:89291870; PMID:2544587

A;Accession: A34274

A;Molecule type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.

J. Cell Biol. 108, 729-736, 1989

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>

A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-397 <KOB>

A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276; PMID:3461443

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-374,'RC' <DIX>

A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992

18

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:00:05 ; Search time 5.4978 Seconds  
(without alignments)  
113.653 Million cell updates/sec

Title: US-10-030-735-18  
Perfect score: 61  
Sequence: 1 FQGVQLQNLRFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	58	95.1	1170	1	TSP1_BOVIN	Q28178 bos taurus
2	58	95.1	1170	1	TSP1_HUMAN	P07996 homo sapien
3	58	95.1	1170	1	TSP1_MOUSE	P35441 mus musculu
4	58	95.1	1173	1	TSP1_XENLA	P35448 xenopus lae
5	41	67.2	467	1	FLII_BUCAI	P57178 buchnera ap
6	40	65.6	1172	1	TSP2_HUMAN	P35442 homo sapien
7	40	65.6	1172	1	TSP2_MOUSE	Q03350 mus musculu
8	40	65.6	1178	1	TSP2_CHICK	P35440 gallus gall
9	38	62.3	212	1	HIS5_NEIMA	Q9jvh3 neisseria m
10	38	62.3	1170	1	TSP2_BOVIN	Q95116 bos taurus
11	38	62.3	1312	1	DPOL_PYRSD	Q51334 pyrococcus
12	38	62.3	1671	1	DPOL_PYRKO	P77933 pyrococcus
13	38	62.3	1699	1	DPOL_THEG8	Q9hh84 thermococcu
14	36	59.0	199	1	HIS5_HAEIN	P44340 haemophilus
15	36	59.0	385	1	YFD0_YEAST	P43567 saccharomyc
16	36	59.0	492	1	CPAC_MOUSE	P56593 mus musculu
17	36	59.0	749	1	NEP_HUMAN	P08473 homo sapien

18	36	59.0	749	1	NEP_MOUSE	Q61391	mus musculus
19	36	59.0	749	1	NEP_RABIT	P08049	oryctolagus
20	36	59.0	749	1	NEP_RAT	P07861	rattus norv
21	36	59.0	1702	1	DPOL_THELI	P30317	thermococcu
22	35	57.4	380	1	PEX2_YARLI	Q99155	yarrowia li
23	35	57.4	392	1	OXDD_BACSU	O34767	bacillus su
24	35	57.4	531	1	MGLC_TREPA	Q57321	treponema p
25	35	57.4	700	1	V018_FOWPV	Q9j5i3	fowlpox vir
26	35	57.4	1349	1	NOS_DROME	Q27571	drosophila
27	35	57.4	1829	1	DPOL_THEST	O33845	thermococcu
28	34	55.7	243	1	HIS4_HELHP	Q7vhy5	helicobacte
29	34	55.7	312	1	OTCC_MYCCC	P59779	mycoplasma
30	34	55.7	562	1	ATKA_YERPE	Q8zd96	yersinia pe
31	34	55.7	585	1	YK05_CAEEL	P34293	caenorhabdi
32	34	55.7	1288	1	OPLA_MOUSE	Q8k010	mus musculus
33	34	55.7	1288	1	OPLA_RAT	P97608	rattus norv
34	33	54.1	196	1	BIM_MOUSE	O54918	mus musculus
35	33	54.1	196	1	BIM_RAT	O88498	rattus norv
36	33	54.1	222	1	PLL2_MOUSE	P09586	mus musculus
37	33	54.1	250	1	UBIE_COXBU	Q83a90	coxiella bu
38	33	54.1	352	1	IDI2_PYRAE	Q8zyf6	pyrobaculum
39	33	54.1	453	1	YEGQ_ECOLI	P76403	escherichia
40	33	54.1	466	1	FLII_BUCAP	Q8ka42	buchnera ap
41	33	54.1	474	1	NU4M_PARTE	P15581	paramecium
42	33	54.1	492	1	CPA1_RAT	P11711	rattus norv
43	33	54.1	492	1	CPA2_RAT	P15149	rattus norv
44	33	54.1	573	1	MLO4_ARATH	O23693	arabidopsis
45	33	54.1	632	1	KU70_CHICK	O93257	gallus gall
46	33	54.1	672	1	HYFB_ECOLI	P23482	escherichia
47	33	54.1	680	1	DNK3_SYNEL	Q8dh10	synechococc
48	33	54.1	764	1	SYFB_HELPY	P56145	helicobacte
49	33	54.1	910	1	SYL_THEAC	Q9hk31	thermoplasm
50	33	54.1	1020	1	ACA1_ARATH	Q37145	arabidopsis
51	33	54.1	1069	1	ACAA_ARATH	Q9szr1	arabidopsis
52	33	54.1	1079	1	APCE_FREDI	P16566	fremyella d
53	33	54.1	1131	1	APCE_ANASP	P80559	anabaena sp
54	33	54.1	1177	1	Y307_MYCGE	P47549	mycoplasma
55	33	54.1	1331	1	XDH_BOVIN	P80457	bos taurus
56	33	54.1	1332	1	XDH_HUMAN	P47989	homo sapien
57	33	54.1	1839	1	ALE3_AZOVI	Q44496	azotobacter
58	33	54.1	3587	1	SRF2_BACSU	Q04747	bacillus su
59	32.5	53.3	308	1	FTSQ_BARBA	Q9x5h9	bartonella
60	32	52.5	160	1	COTF_BACSU	P23261	bacillus su
61	32	52.5	259	1	TRMB_AZOSE	Q8g9c6	azoarcus sp
62	32	52.5	276	1	FLJJ_CAUCR	P02969	caulobacter
63	32	52.5	295	1	TRPC_SYNY3	Q55508	synechocyst
64	32	52.5	315	1	HLGC_STAAU	Q07227	staphylococ
65	32	52.5	315	1	LUKS_STAAU	P31716	staphylococ
66	32	52.5	355	1	GBA2_NEUCR	Q05424	neurospora
67	32	52.5	369	1	YL05_SULSO	P95879	sulfolobus
68	32	52.5	373	1	SI7B_MOUSE	P70277	mus musculus
69	32	52.5	405	1	GLC1_VIBCH	Q9krb5	vibrio chol
70	32	52.5	406	1	T23O_HUMAN	P48775	homo sapien
71	32	52.5	440	1	KRF1_SPVKA	P32216	swinepox vi
72	32	52.5	448	1	Y314_MYCPN	Q50363	mycoplasma
73	32	52.5	457	1	GAL8_KLULA	Q06433	kluyveromyc
74	32	52.5	504	1	COBQ_METTH	O26880	methanobact

75	32	52.5	571	1	ATKA_MYCTU	P96371	mycobacteri
76	32	52.5	620	1	Y512_CHLPN	Q9z840	chlamydia p
77	32	52.5	864	1	MUTS_NEIMA	Q9jw7	neisseria m
78	32	52.5	864	1	MUTS_NEIMB	Q9jx94	neisseria m
79	32	52.5	1014	1	ACA2_ARATH	081108	arabidopsis
80	32	52.5	1015	1	ACA7_ARATH	064806	arabidopsis
81	32	52.5	1045	1	RAG1_XENLA	Q91829	xenopus lae
82	32	52.5	1188	1	S3B1_SCHPO	Q10178	schizosacch
83	32	52.5	1304	1	S3B1_HUMAN	075533	homo sapien
84	32	52.5	1304	1	S3B1_MOUSE	Q99nb9	mus musculu
85	32	52.5	1307	1	S3B1_XENLA	057683	xenopus lae
86	32	52.5	1545	1	ACC9_RAT	Q63563	rattus norv
87	32	52.5	1549	1	ACC9_RABIT	P82451	oryctolagus
88	32	52.5	1668	1	DPOL_THEHY	Q9hh05	thermococcu
89	32	52.5	3712	1	LMA_DROME	Q00174	drosophila
90	31.5	51.6	216	1	CTRD_NEIMA	P57013	neisseria m
91	31.5	51.6	216	1	CTRD_NEIMB	P32016	neisseria m
92	31.5	51.6	217	1	BEXA_HAEIN	P10640	haemophilus
93	31	50.8	102	1	RT14_PARTE	P15759	paramecium
94	31	50.8	126	1	Y334_BUCBP	Q89ag0	buchnera ap
95	31	50.8	195	1	KGUA_MYCGA	Q9kx62	mycoplasma
96	31	50.8	214	1	KGUA_PSESM	Q88be2	pseudomonas
97	31	50.8	249	1	YG61_FUSNN	Q8rie0	fusobacteri
98	31	50.8	275	1	UL34_HSV11	P10218	herpes simp
99	31	50.8	287	1	ATPG_ECOLI	P00837	escherichia
100	31	50.8	288	1	ATPG_VIBAL	P12990	vibrio algi
101	31	50.8	295	1	LIPA_AERPE	Q9y9e3	aeropyrum p
102	31	50.8	295	1	SUOT_MOUSE	P49891	mus musculu
103	31	50.8	326	1	MOAA_YERPE	Q8zgw5	yersinia pe
104	31	50.8	335	1	Y286_STAEP	Q8ctu5	staphylococ
105	31	50.8	335	1	Y479_STAAW	Q8nxy9	staphylococ
106	31	50.8	335	1	Y524_STAAM	Q99w79	staphylococ
107	31	50.8	336	1	SELD_AQUAE	067139	aquifex aeo
108	31	50.8	338	1	PHND_ECOLI	P16682	escherichia
109	31	50.8	353	1	GBQ_CANFA	Q28294	canis famil
110	31	50.8	353	1	GBQ_HOMAM	P91950	homarus ame
111	31	50.8	353	1	GBQ_HUMAN	P50148	homo sapien
112	31	50.8	353	1	GBQ_LYMST	P38411	lymnaea sta
113	31	50.8	353	1	GBQ_MOUSE	P21279	mus musculu
114	31	50.8	353	1	GBQ_PATYE	015975	patinopecte
115	31	50.8	353	1	GBQ_RAT	P82471	rattus norv
116	31	50.8	353	1	GBQ_XENLA	P38410	xenopus lae
117	31	50.8	354	1	GB14_XENLA	073819	xenopus lae
118	31	50.8	354	1	GBQ_LOLFO	P38412	loligo forb
119	31	50.8	355	1	GB14_BOVIN	P38408	bos taurus
120	31	50.8	355	1	GB14_MOUSE	P30677	mus musculu
121	31	50.8	359	1	GB11_BOVIN	P38409	bos taurus
122	31	50.8	359	1	GB11_HUMAN	P29992	homo sapien
123	31	50.8	359	1	GB11_MELGA	P45645	meleagris g
124	31	50.8	359	1	GB11_MOUSE	P21278	mus musculu
125	31	50.8	359	1	GB11_RAT	Q9jid2	rattus norv
126	31	50.8	359	1	GB11_XENLA	P43444	xenopus lae
127	31	50.8	378	1	GB12_MOUSE	P27600	mus musculu
128	31	50.8	378	1	GB12_RAT	Q63210	rattus norv
129	31	50.8	378	1	YIBH_ECOLI	P32107	escherichia
130	31	50.8	405	1	KRF1_VACCP	P29884	vaccinia vi
131	31	50.8	410	1	PGK_PYRFU	P50316	pyrococcus



132	31	50.8	415	1	YBDG_ECOLI	P39455	escherichia
133	31	50.8	439	1	KRF1_VACCC	P21095	vaccinia vi
134	31	50.8	439	1	KRF1_VARV	P33801	variola vir
135	31	50.8	460	1	YEGQ_HAEIN	P44700	haemophilus
136	31	50.8	470	1	DAX1_HUMAN	P51843	homo sapien
137	31	50.8	471	1	DAX1_PIG	P79386	sus scrofa
138	31	50.8	472	1	DAX1_MOUSE	Q61066	mus musculu
139	31	50.8	472	1	DAX1_RAT	P70503	rattus norv
140	31	50.8	494	1	AMY1_SACFI	P21567	saccharomyc
141	31	50.8	504	1	CP7A_CRIGR	P46634	cricetulus
142	31	50.8	512	1	UGTB_CAEEL	Q22180	caenorhabdi
143	31	50.8	513	1	C942_VICSA	P98188	vicia sativ
144	31	50.8	516	1	SR54_STRMU	Q54431	streptococc
145	31	50.8	522	1	C5P3_ARATH	Q9fg23	arabidopsis
146	31	50.8	529	1	RAI2_MOUSE	Q9qvy8	mus musculu
147	31	50.8	530	1	RAI2_HUMAN	Q9y5p3	homo sapien
148	31	50.8	565	1	HEMA_IATKP	P03456	influenza a
149	31	50.8	604	1	SR68_DROME	Q9vss2	drosophila
150	31	50.8	608	1	KU70_MOUSE	P23475	mus musculu

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.  
AC Q28178; Q28179;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Thrombospondin 1 precursor.  
GN THBS1 OR TSP1 OR TSP-1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Holstein; TISSUE=Tooth;  
RX MEDLINE=98173773; PubMed=9507054;  
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,  
RA Inoue H.;  
RT "cDNA cloning of bovine thrombospondin 1 and its expression in  
RT odontoblasts and predentin.";  
RL Biochim. Biophys. Acta 1382:17-22(1998).  
RN [2]  
RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.  
RC TISSUE=Aortic endothelium;  
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-  
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis  
CC and/or maintenance of dentin and dental pulp.

```

CC  -!- SUBUNIT: Homotrimer; disulfide-linked.
CC  -!- TISSUE SPECIFICITY: Odontoblasts.
CC  -!- SIMILARITY: Belongs to the thrombospondin family.
CC  -!- SIMILARITY: Contains 1 VWFC domain.
CC  -!- SIMILARITY: Contains 3 EGF-like domains.
CC  -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC  -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC  -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AB005287; BAA21115.1; -.
DR  EMBL; X87618; CAA60950.1; -.
DR  EMBL; X87619; CAA60951.1; -.
DR  PIR; S55501; S55501.
DR  GlycoSuiteDB; Q28178; -.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR006209; EGF_like.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000884; TSP1.
DR  InterPro; IPR008085; TSP_1.
DR  InterPro; IPR003367; tsp_3.
DR  InterPro; IPR008859; TSPC.
DR  InterPro; IPR003129; TSPN.
DR  InterPro; IPR001007; VWF_C.
DR  Pfam; PF00008; EGF; 2.
DR  Pfam; PF00090; tsp_1; 3.
DR  Pfam; PF02412; tsp_3; 13.
DR  Pfam; PF05735; TSPC; 1.
DR  Pfam; PF02210; TSPN; 1.
DR  Pfam; PF00093; vwc; 1.
DR  PRINTS; PR01705; TSP1REPEAT.
DR  SMART; SM00181; EGF; 3.
DR  SMART; SM00209; TSP1; 3.
DR  SMART; SM00210; TSPN; 1.
DR  SMART; SM00214; VWC; 1.
DR  PROSITE; PS00022; EGF_1; FALSE_NEG.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS50026; EGF_3; 2.
DR  PROSITE; PS50092; TSP1; 3.
DR  PROSITE; PS01208; VWFC_1; 1.
DR  PROSITE; PS50184; VWFC_2; 1.
KW  Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW  EGF-like domain; Signal.
FT      SIGNAL          1      18      BY SIMILARITY.
FT      CHAIN           19     1170     THROMBOSPONDIN 1.
FT      DOMAIN          19     232     HEPARIN-BINDING (POTENTIAL).
FT      DOMAIN          24     221     TSP N-TERMINAL.
FT      DOMAIN          316    373     VWFC.
FT      DOMAIN          379    429     TSP TYPE-1 1.
FT      DOMAIN          435    490     TSP TYPE-1 2.

```

10/030,735 SEQ ID NO:18

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:00:35 ; Search time 31.5595 Seconds  
(without alignments)  
119.971 Million cell updates/sec

Title: US-10-030-735-18  
Perfect score: 61  
Sequence: 1 FQGV LQNLR FVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp Vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%
No.	Score	Match Length DB ID
Description		
-----		

1	58	95.1	229	6	Q28194	Q28194 bos taurus
2	58	95.1	496	13	Q7SY84	Q7sy84 xenopus lae
3	58	95.1	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	58	95.1	1171	11	Q80YQ1	Q80yq1 mus musculu
5	41	67.2	145	5	Q93511	Q93511 caenorhabdi
6	40	65.6	1172	11	Q8CG21	Q8cg21 mus musculu
7	40	65.6	1172	11	Q7TMT3	Q7tmt3 mus musculu
8	39	63.9	727	17	Q8PRY3	Q8pry3 methanosarc
9	39	63.9	733	17	Q8TLX6	Q8tlx6 methanosarc
10	38	62.3	278	5	Q9W251	Q9w251 drosophila
11	38	62.3	593	2	Q8GH66	Q8gh66 mycobacteri
12	38	62.3	750	11	Q8BVV4	Q8bvv4 mus musculu
13	38	62.3	759	16	Q8ZCG3	Q8zcg3 yersinia pe
14	37	60.7	101	1	Q8NKU9	Q8nku9 acidianus a
15	37	60.7	101	16	Q7VC17	Q7vc17 prochloroco
16	37	60.7	217	16	Q8RC20	Q8rc20 thermoanaer
17	37	60.7	258	6	Q9BEZ5	Q9bez5 choloepus d
18	37	60.7	258	6	Q9BEZ6	Q9bez6 choloepus h
19	37	60.7	258	6	Q9BEZ4	Q9bez4 euphractus
20	37	60.7	337	16	Q8CZC1	Q8czc1 streptococc
21	37	60.7	343	3	Q9P866	Q9p866 candida alb
22	37	60.7	471	13	Q90577	Q90577 gallus gall
23	37	60.7	770	13	O93394	O93394 perca flave
24	36	59.0	223	10	Q9LMB5	Q9lmb5 arabidopsis
25	36	59.0	269	16	Q8A780	Q8a780 bacteroides
26	36	59.0	346	16	Q8FFZ0	Q8ffz0 escherichia
27	36	59.0	405	17	Q8U2Q0	Q8u2q0 pyrococcus
28	36	59.0	492	11	Q8VCW9	Q8vcw9 mus musculu
29	36	59.0	493	11	Q9R0R6	Q9r0r6 cricetulus
30	36	59.0	511	16	P73413	P73413 synechocyst
31	36	59.0	595	4	Q9UEU4	Q9ueu4 homo sapien
32	36	59.0	595	5	Q86NR6	Q86nr6 drosophila
33	36	59.0	721	5	Q9VTH0	Q9vth0 drosophila
34	36	59.0	727	10	Q9FZI5	Q9fzi5 arabidopsis
35	36	59.0	750	11	Q8K251	Q8k251 mus musculu
36	36	59.0	876	10	Q9C622	Q9c622 arabidopsis
37	36	59.0	956	13	Q8QG47	Q8qg47 aegithalos
38	36	59.0	1034	5	Q7Z2B9	Q7z2b9 trypanosoma
39	36	59.0	1038	5	O60983	O60983 dictyosteli
40	36	59.0	1354	11	Q9EPW8	Q9epw8 mus musculu
41	36	59.0	1457	12	Q9DY97	Q9dy97 porcine rep
42	36	59.0	1457	12	Q9WJB3	Q9wjb3 porcine rep
43	36	59.0	1457	12	Q9WBQ4	Q9wbq4 porcine rep
44	36	59.0	1457	12	Q91F53	Q91f53 porcine rep
45	36	59.0	1457	12	Q80KX0	Q80kx0 porcine rep
46	36	59.0	1460	12	Q8QQW9	Q8qqw9 porcine rep
47	36	59.0	1460	12	Q8B911	Q8b911 porcine rep
48	36	59.0	1463	12	Q9YN01	Q9yn01 porcine rep
49	36	59.0	1463	12	Q9ENK5	Q9enk5 porcine rep
50	36	59.0	1463	12	Q9E8M9	Q9e8m9 porcine rep
51	36	59.0	1463	12	Q99AV5	Q99av5 porcine rep
52	36	59.0	1463	12	Q99BU5	Q99bu5 porcine rep
53	36	59.0	1463	12	Q7TF56	Q7tf56 porcine rep
54	36	59.0	1480	11	Q80TM9	Q80tm9 mus musculu
55	36	59.0	1504	4	Q9UES6	Q9ues6 homo sapien
56	36	59.0	1504	4	Q7Z2X6	Q7z2x6 homo sapien
57	36	59.0	1528	4	Q9Y2I1	Q9y2i1 homo sapien

58	36	59.0	3956	12	Q9DLN9	Q9dln9 porcine rep
59	36	59.0	3960	12	Q9DLP1	Q9dlp1 porcine rep
60	36	59.0	3960	12	Q9DLN8	Q9dln8 porcine rep
61	36	59.0	3960	12	Q9DLP0	Q9dlp0 porcine rep
62	35	57.4	79	2	P82571	P82571 streptococc
63	35	57.4	92	11	P70317	P70317 mus musculu
64	35	57.4	100	16	Q83E24	Q83e24 coxiella bu
65	35	57.4	146	17	Q976Q1	Q976q1 sulfolobus
66	35	57.4	210	16	Q87ZW0	Q87zw0 pseudomonas
67	35	57.4	246	2	Q9RGX1	Q9rgx1 streptococc
68	35	57.4	247	16	O25286	O25286 helicobacte
69	35	57.4	247	16	Q9ZLS0	Q9zls0 helicobacte
70	35	57.4	301	16	Q7V218	Q7v218 prochloroco
71	35	57.4	394	6	Q8HZW7	Q8hzw7 ornithorhyn
72	35	57.4	403	10	Q9FLG6	Q9flg6 arabidopsis
73	35	57.4	431	5	O17199	O17199 caenorhabdi
74	35	57.4	455	2	Q8KQL8	Q8kql8 streptococc
75	35	57.4	455	16	Q9HZE7	Q9hze7 pseudomonas
76	35	57.4	456	16	Q99ZN6	Q99zn6 streptococc
77	35	57.4	456	16	Q8K7H3	Q8k7h3 streptococc
78	35	57.4	456	16	Q8P117	Q8p117 streptococc
79	35	57.4	459	16	Q9CMS7	Q9cms7 pasteurella
80	35	57.4	465	16	Q8ZCV8	Q8zcv8 yersinia pe
81	35	57.4	586	16	Q819G8	Q819g8 bacillus ce
82	35	57.4	627	16	Q87GG1	Q87gg1 vibrio para
83	35	57.4	834	3	O43048	O43048 schizosacch
84	35	57.4	849	5	Q9V780	Q9v780 drosophila
85	35	57.4	893	16	Q9S258	Q9s258 streptomyce
86	35	57.4	951	13	Q8QFY3	Q8qfy3 vanga curvi
87	35	57.4	952	13	Q8QFY5	Q8qfy5 troglodytes
88	35	57.4	952	13	Q8QFY9	Q8qfy9 thamnophilu
89	35	57.4	953	13	Q8QG14	Q8qg14 nectarinia
90	35	57.4	955	13	Q8QG08	Q8qg08 pardalotus
91	35	57.4	956	13	Q8QFY7	Q8qfy7 toxorhamphu
92	35	57.4	956	13	Q9W6P9	Q9w6p9 tyrannus ty
93	35	57.4	957	13	Q8QFZ5	Q8qfz5 regulus cal
94	35	57.4	957	13	Q8QG23	Q8qg23 lanius ludo
95	35	57.4	957	13	Q8QFZ4	Q8qfz4 rupicola ru
96	35	57.4	957	13	Q8QG38	Q8qg38 cinclus cin
97	35	57.4	957	13	Q8QG12	Q8qg12 oriolus lar
98	35	57.4	957	13	Q8QG37	Q8qg37 cisticola a
99	35	57.4	957	13	Q8QG10	Q8qg10 pachycephal
100	35	57.4	957	13	Q8QG19	Q8qg19 menura nova
101	35	57.4	957	13	Q8QFZ1	Q8qfz1 sturnus vul
102	35	57.4	957	13	Q8QG04	Q8qg04 picathartes
103	35	57.4	957	13	Q8QG46	Q8qg46 aegithina t
104	35	57.4	957	13	Q8QG15	Q8qg15 muscicapa s
105	35	57.4	957	13	Q8QG26	Q8qg26 hirundo pyr
106	35	57.4	957	13	Q8QG28	Q8qg28 furnarius r
107	35	57.4	957	13	Q8QG24	Q8qg24 irena cyano
108	35	57.4	957	13	Q8QG21	Q8qg21 melanochari
109	35	57.4	957	13	Q8QG27	Q8qg27 garrulax mi
110	35	57.4	957	13	Q8QG43	Q8qg43 artamus leu
111	35	57.4	957	13	Q8QFZ0	Q8qfz0 sylvia nana
112	35	57.4	957	13	Q8QG13	Q8qg13 oedistoma i
113	35	57.4	957	13	Q8QG30	Q8qg30 formicarius
114	35	57.4	957	13	Q8QG18	Q8qg18 mimus patag

115	35	57.4	957	13	Q8QG03	Q8qg03 pipra coron
116	35	57.4	957	13	Q8QG11	Q8qg11 orthonyx sp
117	35	57.4	957	13	Q8QFY1	Q8qfy1 zosterops s
118	35	57.4	957	13	Q8QG17	Q8qg17 monarcha ax
119	35	57.4	957	13	Q8QG45	Q8qg45 alauda arve
120	35	57.4	957	13	Q8QG35	Q8qg35 coracina li
121	35	57.4	961	13	Q8QG07	Q8qg07 parula amer
122	35	57.4	961	13	Q8QFY8	Q8qfy8 thraupis cy
123	35	57.4	961	13	Q8QG41	Q8qg41 cardinalis
124	35	57.4	961	13	Q8QG31	Q8qg31 emberiza sc
125	35	57.4	961	13	Q9W6P8	Q9w6p8 passer mont
126	35	57.4	962	5	Q7YZ58	Q7yz58 cryptospori
127	35	57.4	1054	17	O29238	O29238 archaeoglob
128	35	57.4	1349	5	Q7YU33	Q7yu33 drosophila
129	35	57.4	3242	9	Q859P9	Q859p9 bacterioph
130	34.5	56.6	79	17	O27334	O27334 methanobact
131	34	55.7	86	16	Q7UYA0	Q7uya0 rhodopirell
132	34	55.7	132	2	Q9EZM6	Q9ezm6 staphylococ
133	34	55.7	133	16	Q99T48	Q99t48 staphylococ
134	34	55.7	151	16	Q7UDW7	Q7udw7 rhodopirell
135	34	55.7	163	16	Q892M4	Q892m4 clostridium
136	34	55.7	172	16	Q81FM9	Q81fm9 bacillus ce
137	34	55.7	175	10	Q84RF8	Q84rf8 gossypium b
138	34	55.7	193	16	Q8ERG7	Q8erg7 oceanobacil
139	34	55.7	208	5	Q8MSE6	Q8mse6 drosophila
140	34	55.7	210	2	Q9ZEN6	Q9zen6 wolinellet s
141	34	55.7	210	11	Q9D9C0	Q9d9c0 mus musculu
142	34	55.7	219	16	Q8ENF8	Q8enf8 oceanobacil
143	34	55.7	241	5	Q9VDR0	Q9vdr0 drosophila
144	34	55.7	243	11	Q99NN3	Q99nn3 dipodomys h
145	34	55.7	243	16	Q7VHY5	Q7vhy5 helicobacte
146	34	55.7	258	6	Q9BEX1	Q9bex1 nycteris th
147	34	55.7	262	6	Q8WN46	Q8wn46 nycteris th
148	34	55.7	265	16	P70972	P70972 bacillus su
149	34	55.7	290	16	Q8YZP4	Q8yzp4 anabaena sp
150	34	55.7	301	16	Q8A3M4	Q8a3m4 bacteroides

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.  
AC Q28194;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thrombospondin-1 (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 95.1%; Score 58; DB 6; Length 229;  
 Best Local Similarity 91.7%; Pred. No. 0.0033;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQLNLRVFV 12  
 |||||:||||  
 Db 190 FQGVQLNVRVFV 201

## RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

-2004 CompuGen Ltd.

OM protein -protein search, using sw model

Run on: April 7, 2004, 17:59:29 ; Search time 48.3172 Seconds  
(without alignments)  
70.173 Million cell updates/sec

Title: US-10-030-735-19  
Perfect score: 61  
Sequence: 1 FQGV LQDVR FVF 12

Scoring table: BLOSUM 62  
Gapop 10.0, Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database: A\_Geneseq\_29Jan04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred.No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

%  
Result Query



No.	Score	Match	Length	DB	ID	Description
1	61	100.0	12	4	AAB35380	Aab35380 Alpha3bet
2	56	91.8	12	4	AAB35352	Aab35352 Alpha3bet
3	56	91.8	12	4	AAB35378	Aab35378 Alpha3bet
4	56	91.8	12	6	ABG72834	Abg72834 Thrombosp
5	56	91.8	459	4	AAU02916	Aau02916 Angiotens
6	56	91.8	466	3	AAB43602	Aab43602 Human can
7	56	91.8	546	4	AAU02915	Aau02915 Angiotens
8	56	91.8	555	4	AAU02914	Aau02914 Angiotens
9	56	91.8	731	4	AAU02913	Aau02913 Angiotens
10	56	91.8	1152	3	AAB00042	Aab00042 Human thr
11	56	91.8	1152	5	AAU74771	Aau74771 Human thr
12	56	91.8	1152	5	ABB82285	Abb82285 Human thr
13	56	91.8	1170	4	AAB74450	Aab74450 Human var
14	56	91.8	1170	4	AAB90800	Aab90800 Human she
15	56	91.8	1170	5	AAE25030	Aae25030 Human thr
16	56	91.8	1170	5	AAU75315	Aau75315 Human thr
17	56	91.8	1170	6	ABP96780	Abp96780 Human COP
18	56	91.8	1170	6	ABU03474	Abu03474 Angiogene
19	56	91.8	1170	6	ABG74673	Abg74673 Human THB
20	56	91.8	1170	6	AAE36228	Aae36228 Human THB
21	56	91.8	1170	7	ABR62059	Abr62059 Human thr
22	55	90.2	12	4	AAB35379	Aab35379 Alpha3bet
23	55	90.2	12	4	AAB35376	Aab35376 Alpha3bet
24	53	86.9	12	4	AAB35373	Aab35373 Alpha3bet
25	53	86.9	12	4	AAB35361	Aab35361 Alpha3bet
26	53	86.9	12	4	AAB35381	Aab35381 Alpha3bet
27	52	85.2	12	4	AAB35364	Aab35364 Alpha3bet
28	52	85.2	12	4	AAB35374	Aab35374 Alpha3bet
29	51	83.6	12	4	AAB35371	Aab35371 Alpha3bet
30	51	83.6	12	4	AAB35368	Aab35368 Alpha3bet
31	51	83.6	12	4	AAB35367	Aab35367 Alpha3bet
32	50	82.0	11	4	AAB35360	Aab35360 Alpha3bet
33	50	82.0	11	4	AAB35357	Aab35357 Alpha3bet
34	50	82.0	12	4	AAB35366	Aab35366 Alpha3bet
35	50	82.0	12	4	AAB35370	Aab35370 Alpha3bet
36	50	82.0	12	4	AAB35362	Aab35362 Alpha3bet
37	50	82.0	12	4	AAB35363	Aab35363 Alpha3bet
38	50	82.0	12	4	AAB35369	Aab35369 Alpha3bet
39	48	78.7	12	4	AAB35365	Aab35365 Alpha3bet
40	46	75.4	10	4	AAB35355	Aab35355 Alpha3bet
41	46	75.4	12	4	AAB35347	Aab35347 Alpha3bet
42	40	65.6	9	4	AAB35375	Aab35375 Alpha3bet
43	40	65.6	15	6	ABR75751	Abr75751 Liver res
44	39	63.9	9	4	AAB35359	Aab35359 Alpha3bet

45	39	63.9	226	6	ABP77949	Abp77949 N. gonorr
46	38	62.3	16	3	AAB19679	Aab19679 Human thr
47	38	62.3	151	3	AAB53711	Aab53711 Human col
48	38	62.3	1172	3	AAB19677	Aab19677 Human thr
49	38	62.3	1172	3	AAB00043	Aab00043 Human thr
50	38	62.3	1172	5	AAU76902	Aau76902 Human Thr
51	38	62.3	1172	5	AAU74786	Aau74786 Mouse thr
52	38	62.3	1172	5	AAU74788	Aau74788 Human thr
53	38	62.3	1172	5	AAE25031	Aae25031 Human thr
54	38	62.3	1172	5	ABB72334	Abb72334 Rat prote
55	38	62.3	1172	5	ABP68622	Abp68622 Human pan
56	38	62.3	1172	6	ABG72839	Abg72839 Human ang
57	38	62.3	1172	6	ABG74674	Abg74674 Human THB
58	38	62.3	1172	6	AAE36411	Aae36411 Human THB
59	38	62.3	1172	7	ADD31095	Add31095 Human thr
60	37	60.7	233	6	ABM68087	Abm68087 Photorhab
61	37	60.7	258	4	AAG79221	Aag79221 Amino aci
62	37	60.7	810	2	AAW37500	Aaw37500 Human nel
63	37	60.7	810	6	ABR58704	Abr58704 Human can
64	37	60.7	810	6	ABR56521	Abr56521 Human nel
65	37	60.7	810	7	ADE62649	Ade62649 Human Pro
66	37	60.7	810	7	ADE62645	Ade62645 Human Pro
67	37	60.7	810	7	ADE62643	Ade62643 Rat Prote
68	37	60.7	810	7	ADE62647	Ade62647 Rat Prote
69	37	60.7	1039	4	ABB63124	Abb63124 Drosophil
70	36	59.0	162	7	ADC88274	Adc88274 Ribosomal
71	36	59.0	163	4	AAU35478	Aau35478 Haemophil
72	36	59.0	163	6	ABU30324	Abu30324 Protein e
73	36	59.0	163	6	ABU39469	Abu39469 Protein e
74	36	59.0	421	6	ABU34028	Abu34028 Protein e
75	36	59.0	459	6	ABU39156	Abu39156 Protein e
76	36	59.0	637	4	ABB71200	Abb71200 Drosophil
77	36	59.0	920	6	ABU26032	Abu26032 Protein e
78	35	57.4	8	4	AAB35358	Aab35358 Alpha3bet
79	35	57.4	48	2	AAR06749	Aar06749 Peptide #
80	35	57.4	136	4	AAB87620	Aab87620 Bovine ma
81	35	57.4	192	2	AAR06895	Aar06895 Metallopr
82	35	57.4	192	6	ABU06411	Abu06411 Maize SS1
83	35	57.4	213	6	ABU06592	Abu06592 Maize SSI
84	35	57.4	214	6	ABU06793	Abu06793 Maize SSI
85	35	57.4	232	5	ABB92454	Abb92454 Herbicida
86	35	57.4	421	6	ABU36094	Abu36094 Protein e
87	35	57.4	421	7	ADB74281	Adb74281 Mycobacte
88	35	57.4	446	6	ABM71622	Abm71622 Staphyloc
89	35	57.4	682	6	AAO16797	Aao16797 Human sec
90	35	57.4	816	2	AAW37501	Aaw37501 Human nel

91	35	57.4	816	6	ABR56522	Abr56522 Human nel
92	35	57.4	944	6	ABP96125	Abp96125 Phoma bet
93	35	57.4	2146	4	ABB62317	Abb62317 Drosophil
94	35	57.4	3606	4	ABB62595	Abb62595 Drosophil
95	34	55.7	8	4	AAB35356	Aab35356 Alpha3bet
96	34	55.7	40	4	AAM15164	Aam15164 Peptide #
97	34	55.7	40	4	ABB34156	Abb34156 Peptide #
98	34	55.7	40	4	AAM27621	Aam27621 Peptide #
99	34	55.7	40	4	ABB28990	Abb28990 Peptide #
100	34	55.7	40	4	ABB19601	Abb19601 Protein #
101	34	55.7	40	4	AAM67330	Aam67330 Human bon
102	34	55.7	40	4	AAM54950	Aam54950 Human bra
103	34	55.7	40	4	ABG48992	Abg48992 Human liv
104	34	55.7	40	4	AAM02903	Aam02903 Peptide #
105	34	55.7	40	5	ABG36971	Abg36971 Human pep
106	34	55.7	50	4	AAU52387	Aau52387 Propionib
107	34	55.7	50	6	ABM48906	Abm48906 Propionib
108	34	55.7	147	6	ADA22667	Ada22667 A. gossyp
109	34	55.7	158	2	AAW19579	Aaw19579 Mouse ant
110	34	55.7	159	6	ABU02784	Abu02784 S. pneumo
111	34	55.7	159	6	ABU02511	Abu02511 S. pneumo
112	34	55.7	190	4	AAM00921	Aam00921 Human bon
113	34	55.7	195	6	ABU01368	Abu01368 S. pneumo
114	34	55.7	211	4	AAB79825	Aab79825 Corynebac
115	34	55.7	211	4	AAG92040	Aag92040 C glutami
116	34	55.7	229	3	AAG47140	Aag47140 Arabidops
117	34	55.7	229	3	AAG14083	Aag14083 Arabidops
118	34	55.7	269	3	AAG47139	Aag47139 Arabidops
119	34	55.7	269	3	AAG14082	Aag14082 Arabidops
120	34	55.7	307	4	ABG11255	Abg11255 Novel hum
121	34	55.7	318	3	AAG14081	Aag14081 Arabidops
122	34	55.7	318	3	AAG47138	Aag47138 Arabidops
123	34	55.7	336	3	AAY79181	Aay79181 Haematopo
124	34	55.7	374	4	AAG81674	Aag81674 S. epider
125	34	55.7	375	5	ABP40205	Abp40205 Staphyloc
126	34	55.7	403	4	ABG11258	Abg11258 Novel hum
127	34	55.7	440	2	AAW26414	Aaw26414 Swinepox
128	34	55.7	440	4	AAB68235	Aab68235 Protein e
129	34	55.7	445	7	ADE37191	Ade37191 Plant yie
130	34	55.7	481	5	ABB08866	Abb08866 Corynebac
131	34	55.7	483	2	AAR82690	Aar82690 Brevibact
132	34	55.7	483	4	AAB79623	Aab79623 Corynebac
133	34	55.7	483	4	AAG91088	Aag91088 C glutami
134	34	55.7	536	2	AAW01107	Aaw01107 Endonucle
135	34	55.7	539	4	ABG99090	Abg99090 Pyrococcu
136	34	55.7	607	6	ABU48470	Abu48470 Protein e

137	34	55.7	667	4	AAU49002	Aau49002 Propionib
138	34	55.7	667	6	ABM45521	Abm45521 Propionib
139	34	55.7	721	4	ABB62165	Abb62165 Drosophil
140	34	55.7	722	4	AAM93962	Aam93962 Human pol
141	34	55.7	753	2	AAW57747	Aaw57747 Leukocyte
142	34	55.7	753	5	ABB07637	Abb07637 Human spe
143	34	55.7	771	4	ABG11257	Abg11257 Novel hum
144	34	55.7	806	6	ADA22669	Ada22669 A. gossyp
145	34	55.7	885	4	ABG11256	Abg11256 Novel hum
146	34	55.7	1010	7	ADE58837	Ade58837 Human Pro
147	34	55.7	1010	7	ADE58841	Ade58841 Human Pro
148	34	55.7	1019	2	AAR38096	Aar38096 Pyrococcu
149	34	55.7	1019	2	AAR81554	Aar81554 Pyrococcu
150	34	55.7	1039	7	ADE58839	Ade58839 Rat Prote

## ALIGNMENTS

### RESULT 1

AAB35380

ID AAB35380 standard; peptide; 12 AA.

XX

AC AAB35380;

XX

DT 08-MAY-2001 (first entry)

XX

DE Alpha3beta1 integrin binding peptide #45.

XX

KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;

KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;

KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX

OS Synthetic.

XX

PN WO200105812-A2.

XX

PD 25-JAN-2001.

XX

PF 12-JUL-2000; 2000WO-US018986.

XX

PR 15-JUL-1999; 99US-0144549P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Roberts DD, Krutzsch HC;

XX

DR WPI; 2001-182656/18.

XX

PT New peptides that bind to or are recognized by alpha3-beta1 integrins,

PT useful for inhibiting cell adhesion to extracellular matrix, cell

PT motility and proliferation and for treating rheumatoid arthritis and

PT cancer.

XX

PS Claim 4; Page 34; 84pp; English.

XX

CC The present invention provides a number of peptides which bind to

CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion

CC and motility, and in the treatment of cancer, diabetic retinopathy,

CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis

CC and restenosis. The present sequence is an example of one of the peptides

CC of the invention

XX

SQ Sequence 12 AA;

Query Match 100.0%; Score 61; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 0.00015;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQLQDVRFVF 12

|||||||

Db 1 FQGVQLQDVRFVF 12

## RESULT 2

AAB35352

ID AAB35352 standard; peptide; 12 AA.

XX

AC AAB35352;

XX

DT 08-MAY-2001 (first entry)

XX

DE Alpha3beta1 integrin binding peptide #17.

XX

KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;

KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;

KW macular degeneration; psoriasis; cell adhesion; cell motility.

XX

OS Synthetic.

XX

PN WO200105812-A2.

XX

PD 25-JAN-2001.

XX

PF 12-JUL-2000; 2000WO-US018986.

XX

PR 15-JUL-1999; 99US-0144549P.

XX

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Roberts DD, Kruttsch HC;

XX

DR WPI; 2001-182656/18.

XX

PT New peptides that bind to or are recognized by alpha3-beta1 integrins,

PT useful for inhibiting cell adhesion to extracellular matrix, cell

PT motility and proliferation and for treating rheumatoid arthritis and

PT cancer.

XX

PS Claim 4; Page 34; 84pp; English.

XX

CC The present invention provides a number of peptides which bind to

CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion

CC and motility, and in the treatment of cancer, diabetic retinopathy,

CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis

CC and restenosis. The present sequence is an example of one of the peptides

CC of the invention

XX

SQ Sequence 12 AA;

Query Match 91.8%; Score 56; DB 4; Length 12;

Best Local Similarity 91.7%; Pred. No. 0.0013;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLDVRFVF 12

|||||:||||

Db 1 FQGVLQNVRVFVF 12

RESULT 3

AAB35378

ID AAB35378 standard; peptide; 12 AA.

XX

AC AAB35378;

XX

DT 08-MAY-2001 (first entry)

XX

DE Alpha3beta1 integrin binding peptide #43.

XX  
 KW Alpha3beta1 integrin; angiogenesis; cell proliferation; cancer;  
 KW diabetic retinopathy; restenosis; atherosclerosis; rheumatoid arthritis;  
 KW macular degeneration; psoriasis; cell adhesion; cell motility.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200105812-A2.  
 XX  
 PD 25-JAN-2001.  
 XX  
 PF 12-JUL-2000; 2000WO-US018986.  
 XX  
 PR 15-JUL-1999; 99US-0144549P.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Roberts DD, Kruttsch HC;  
 XX  
 DR WPI; 2001-182656/18.  
 XX  
 PT New peptides that bind to or are recognized by alpha3-beta1 integrins,  
 PT useful for inhibiting cell adhesion to extracellular matrix, cell  
 PT motility and proliferation and for treating rheumatoid arthritis and  
 PT cancer.  
 XX  
 PS Example 2; Page 34; 84pp; English.  
 XX  
 CC The present invention provides a number of peptides which bind to  
 CC alpha3beta1 integrins. They are useful in the modulation of cell adhesion  
 CC and motility, and in the treatment of cancer, diabetic retinopathy,  
 CC rheumatoid arthritis, macular degeneration, atherosclerosis, psoriasis  
 CC and restenosis. The present sequence is an example of one of the peptides  
 CC of the invention  
 XX  
 SQ Sequence 12 AA;

Query Match 91.8%; Score 56; DB 4; Length 12;  
 Best Local Similarity 91.7%; Pred. No. 0.0013;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGV LQDV RFVF 12  
 |||||:||||  
 Db 1 FQGV LQNVR FVF 12

RESULT 4

ABG72834

ID ABG72834 standard; peptide; 12 AA.

XX

AC ABG72834;

XX

DT 24-FEB-2003 (first entry)

XX

DE Thrombospondin-1 sequence containing synthetic peptide.

XX

KW Human; thrombospondin-1; cytostatic; immunostimulant; cancer;

KW epithelial cancer; lung cancer; papillary renal cell carcinoma;

KW colon cancer; small-cell lung cancer; SCLC; melanoma.

XX

OS Synthetic.

XX

PN WO200281630-A2.

XX

PD 17-OCT-2002.

XX

PF 03-APR-2002; 2002WO-US010535.

XX

PR 06-APR-2001; 2001US-0281994P.

XX

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX

PI Roberts DD, Krutzsch HC;

XX

DR WPI; 2003-103329/09.

XX

PT A new diagnosis for cancer other than prostate cancer in a mammal useful

PT to detect cancer including lung cancer, particularly small cell lung

PT cancer and melanoma comprises detecting semenogelin in a sample.

XX

PS Example 1; Page 14; 32pp; English.

XX

CC The invention relates to diagnosing cancer other than prostate cancer in

CC a male mammal, comprising assaying a test sample for increased level of

CC semenogelin, or cancer in a female by assaying for the presence of

CC semenogelin. Administering a semenogelin protein or polypeptide fragment

CC or a semenogelin-specific antibody or active fragment, or a recombinant

CC vector expressing the protein or antibody, is useful for inducing an

CC immune response to a cancer in a mammal, where the cancer is not prostate

CC cancer and semenogelin is a marker. The invention is used to diagnose

CC cancer, particularly of epithelial origin such as lung cancer, papillary

CC renal cell carcinoma, colon cancer, especially small-cell lung cancer



CC (SCLC), or a melanoma. The present sequence represents the amino acid  
CC sequence of the thrombospondin-1 sequence containing synthetic peptide  
CC which binds to alpha-3-beta-1 integrin  
XX  
SQ Sequence 12 AA;

Query Match 91.8%; Score 56; DB 6; Length 12;  
Best Local Similarity 91.7%; Pred. No. 0.0013;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGV LQDV RFVF 12  
|||||:||||  
Db 1 FQGV LQNVR FVF 12

SEQ ID NO: 19

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:04:45 ; Search time 12.9515 Seconds  
(without alignments)  
47.833 Million cell updates/sec

Title: US-10-030-735-19  
Perfect score: 61  
Sequence: 1 FQGVLDVRFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	56	91.8	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	39	63.9	432	4	US-09-489-039A-7909	Sequence 7909, Ap
3	38	62.3	1172	1	US-08-313-288B-19	Sequence 19, Appl
4	37	60.7	810	2	US-08-820-170A-34	Sequence 34, Appl
5	37	60.7	810	3	US-09-055-699-34	Sequence 34, Appl
6	37	60.7	810	3	US-09-273-565-34	Sequence 34, Appl
7	37	60.7	810	4	US-09-565-538-34	Sequence 34, Appl
8	37	60.7	810	4	US-09-661-468-34	Sequence 34, Appl
9	37	60.7	810	4	US-09-976-165-34	Sequence 34, Appl
10	36	59.0	162	4	US-09-732-210-527	Sequence 527, App
11	36	59.0	295	3	US-09-150-133-13	Sequence 13, Appl

12	36	59.0	295	3	US-09-150-141-13	Sequence 13, Appl
13	36	59.0	295	3	US-09-374-493-13	Sequence 13, Appl
14	36	59.0	295	3	US-09-374-824-13	Sequence 13, Appl
15	36	59.0	295	3	US-09-374-492-13	Sequence 13, Appl
16	36	59.0	295	4	US-09-785-343-13	Sequence 13, Appl
17	35	57.4	175	4	US-09-252-991A-18826	Sequence 18826, A
18	35	57.4	206	4	US-09-252-991A-20515	Sequence 20515, A
19	35	57.4	421	4	US-08-311-731A-30	Sequence 30, Appl
20	35	57.4	449	4	US-09-489-039A-8009	Sequence 8009, Ap
21	35	57.4	816	2	US-08-820-170A-37	Sequence 37, Appl
22	35	57.4	816	3	US-09-055-699-37	Sequence 37, Appl
23	35	57.4	816	3	US-09-273-565-37	Sequence 37, Appl
24	35	57.4	816	4	US-09-565-538-37	Sequence 37, Appl
25	35	57.4	816	4	US-09-661-468-37	Sequence 37, Appl
26	35	57.4	816	4	US-09-976-165-37	Sequence 37, Appl
27	34	55.7	35	3	US-08-942-686-5	Sequence 5, Appli
28	34	55.7	158	2	US-08-653-402B-6	Sequence 6, Appli
29	34	55.7	169	4	US-09-134-000C-3511	Sequence 3511, Ap
30	34	55.7	273	4	US-09-252-991A-20770	Sequence 20770, A
31	34	55.7	375	4	US-09-134-001C-5050	Sequence 5050, Ap
32	34	55.7	440	1	US-08-307-499-15	Sequence 15, Appl
33	34	55.7	440	3	US-09-299-268-15	Sequence 15, Appl
34	34	55.7	539	4	US-09-800-170-16	Sequence 16, Appl
35	34	55.7	753	3	US-08-942-686-2	Sequence 2, Appli
36	34	55.7	944	4	US-09-134-000C-5578	Sequence 5578, Ap
37	34	55.7	1019	1	US-08-271-364A-7	Sequence 7, Appli
38	34	55.7	1019	2	US-08-222-715B-26	Sequence 26, Appl
39	33	54.1	131	4	US-09-252-991A-22819	Sequence 22819, A
40	33	54.1	156	4	US-09-056-556-227	Sequence 227, App
41	33	54.1	156	4	US-09-072-596-222	Sequence 222, App
42	33	54.1	156	4	US-09-072-967-227	Sequence 227, App
43	33	54.1	187	4	US-09-252-991A-17115	Sequence 17115, A
44	33	54.1	220	4	US-08-134-231C-26	Sequence 26, Appl
45	33	54.1	220	4	US-08-728-160-26	Sequence 26, Appl
46	33	54.1	348	4	US-09-489-039A-13746	Sequence 13746, A
47	33	54.1	408	4	US-09-328-352-6550	Sequence 6550, Ap
48	33	54.1	544	4	US-09-252-991A-30409	Sequence 30409, A
49	33	54.1	596	4	US-09-134-001C-5544	Sequence 5544, Ap
50	33	54.1	659	4	US-09-252-991A-23467	Sequence 23467, A
51	33	54.1	710	4	US-09-252-991A-25700	Sequence 25700, A
52	33	54.1	715	2	US-08-484-993B-10	Sequence 10, Appl
53	33	54.1	715	2	US-08-484-158B-10	Sequence 10, Appl
54	33	54.1	715	2	US-08-484-596A-10	Sequence 10, Appl
55	33	54.1	715	2	US-08-480-150A-10	Sequence 10, Appl
56	33	54.1	715	3	US-08-458-731-10	Sequence 10, Appl
57	33	54.1	715	3	US-08-149-223A-10	Sequence 10, Appl
58	33	54.1	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
59	33	54.1	1242	4	US-09-107-532A-5241	Sequence 5241, Ap
60	33	54.1	1619	4	US-09-392-812A-4	Sequence 4, Appli
61	33	54.1	1704	4	US-09-392-812A-2	Sequence 2, Appli
62	32.5	53.3	563	4	US-09-489-039A-12963	Sequence 12963, A
63	32	52.5	107	4	US-09-198-452A-584	Sequence 584, App
64	32	52.5	132	4	US-09-489-039A-11744	Sequence 11744, A
65	32	52.5	265	4	US-09-252-991A-29034	Sequence 29034, A
66	32	52.5	267	4	US-09-107-532A-4931	Sequence 4931, Ap
67	32	52.5	346	4	US-09-198-452A-60	Sequence 60, Appl
68	32	52.5	358	2	US-08-823-986A-2	Sequence 2, Appli

69	32	52.5	358	3	US-08-784-385-2	Sequence 2, Appli
70	32	52.5	358	5	PCT-US95-10403-2	Sequence 2, Appli
71	32	52.5	368	4	US-09-819-607-5	Sequence 5, Appli
72	32	52.5	390	3	US-08-650-766-7	Sequence 7, Appli
73	32	52.5	390	3	US-08-922-635-6	Sequence 6, Appli
74	32	52.5	390	4	US-09-389-487-7	Sequence 7, Appli
75	32	52.5	420	4	US-09-489-039A-8902	Sequence 8902, Ap
76	32	52.5	497	4	US-09-134-000C-4486	Sequence 4486, Ap
77	32	52.5	550	4	US-09-907-794A-227	Sequence 227, App
78	32	52.5	550	4	US-09-905-125A-227	Sequence 227, App
79	32	52.5	550	4	US-09-902-775A-227	Sequence 227, App
80	32	52.5	559	4	US-09-364-206-47	Sequence 47, Appl
81	32	52.5	596	4	US-09-134-000C-6209	Sequence 6209, Ap
82	32	52.5	651	3	US-08-650-766-6	Sequence 6, Appli
83	32	52.5	651	3	US-08-922-635-5	Sequence 5, Appli
84	32	52.5	651	4	US-09-389-487-6	Sequence 6, Appli
85	32	52.5	667	4	US-09-252-991A-32588	Sequence 32588, A
86	32	52.5	1022	1	US-08-271-364A-8	Sequence 8, Appli
87	32	52.5	1022	2	US-08-222-715B-27	Sequence 27, Appl
88	32	52.5	1070	3	US-08-922-635-22	Sequence 22, Appl
89	32	52.5	1288	3	US-08-762-428A-6	Sequence 6, Appli
90	32	52.5	1504	4	US-09-364-206-2	Sequence 2, Appli
91	31	50.8	129	4	US-09-543-681A-7608	Sequence 7608, Ap
92	31	50.8	229	3	US-09-045-973-7	Sequence 7, Appli
93	31	50.8	242	4	US-09-134-001C-2868	Sequence 2868, Ap
94	31	50.8	249	4	US-09-252-991A-21894	Sequence 21894, A
95	31	50.8	253	4	US-09-489-039A-13477	Sequence 13477, A
96	31	50.8	279	5	PCT-US91-00899-8	Sequence 8, Appli
97	31	50.8	313	4	US-09-540-236-2969	Sequence 2969, Ap
98	31	50.8	369	4	US-09-489-039A-8053	Sequence 8053, Ap
99	31	50.8	371	2	US-08-378-617A-10	Sequence 10, Appl
100	31	50.8	371	2	US-08-378-617A-12	Sequence 12, Appl
101	31	50.8	394	1	US-07-914-281-4	Sequence 4, Appli
102	31	50.8	394	1	US-08-393-246-4	Sequence 4, Appli
103	31	50.8	394	1	US-08-525-058A-4	Sequence 4, Appli
104	31	50.8	394	2	US-08-696-731-4	Sequence 4, Appli
105	31	50.8	394	3	US-09-042-531-4	Sequence 4, Appli
106	31	50.8	394	4	US-08-379-040-2	Sequence 2, Appli
107	31	50.8	394	5	PCT-US91-00899-9	Sequence 9, Appli
108	31	50.8	409	4	US-09-489-039A-14217	Sequence 14217, A
109	31	50.8	420	4	US-09-252-991A-19389	Sequence 19389, A
110	31	50.8	499	4	US-09-107-532A-5524	Sequence 5524, Ap
111	31	50.8	516	4	US-09-291-170A-1	Sequence 1, Appli
112	31	50.8	516	4	US-09-724-884-1	Sequence 1, Appli
113	31	50.8	541	3	US-08-687-590-28	Sequence 28, Appl
114	31	50.8	541	4	US-09-311-311C-25	Sequence 25, Appl
115	31	50.8	605	4	US-09-252-991A-19861	Sequence 19861, A
116	31	50.8	618	4	US-09-134-000C-6538	Sequence 6538, Ap
117	31	50.8	620	3	US-09-232-200-61	Sequence 61, Appl
118	31	50.8	620	4	US-09-232-197-61	Sequence 61, Appl
119	31	50.8	620	4	US-09-232-201-61	Sequence 61, Appl
120	31	50.8	620	4	US-09-232-195-61	Sequence 61, Appl
121	31	50.8	620	4	US-09-418-963-4	Sequence 4, Appli
122	31	50.8	634	4	US-09-252-991A-24131	Sequence 24131, A
123	31	50.8	884	4	US-09-543-681A-5437	Sequence 5437, Ap
124	30	49.2	113	4	US-09-134-001C-5539	Sequence 5539, Ap
125	30	49.2	131	4	US-09-732-210-1647	Sequence 1647, Ap

126	30	49.2	153	2	US-08-387-942C-51	Sequence 51, Appl
127	30	49.2	186	4	US-09-328-352-7876	Sequence 7876, Ap
128	30	49.2	203	4	US-09-252-991A-24219	Sequence 24219, A
129	30	49.2	207	4	US-09-252-991A-31339	Sequence 31339, A
130	30	49.2	218	4	US-08-849-764C-3	Sequence 3, Appli
131	30	49.2	218	4	US-09-262-087-3	Sequence 3, Appli
132	30	49.2	218	4	US-08-463-261B-9	Sequence 9, Appli
133	30	49.2	220	1	US-08-588-163-3	Sequence 3, Appli
134	30	49.2	220	2	US-09-111-070-3	Sequence 3, Appli
135	30	49.2	220	4	US-09-540-530-2	Sequence 2, Appli
136	30	49.2	220	4	US-08-134-231C-27	Sequence 27, Appl
137	30	49.2	220	4	US-09-660-107-11	Sequence 11, Appl
138	30	49.2	220	4	US-08-728-160-27	Sequence 27, Appl
139	30	49.2	226	3	US-09-176-657-3	Sequence 3, Appli
140	30	49.2	226	4	US-09-421-299-3	Sequence 3, Appli
141	30	49.2	265	4	US-09-540-236-3285	Sequence 3285, Ap
142	30	49.2	298	4	US-09-489-039A-8600	Sequence 8600, Ap
143	30	49.2	308	3	US-08-952-089A-13	Sequence 13, Appl
144	30	49.2	330	4	US-09-543-681A-7116	Sequence 7116, Ap
145	30	49.2	341	4	US-09-252-991A-19640	Sequence 19640, A
146	30	49.2	346	4	US-09-134-000C-6034	Sequence 6034, Ap
147	30	49.2	379	4	US-09-491-577-4	Sequence 4, Appli
148	30	49.2	409	4	US-09-328-352-4570	Sequence 4570, Ap
149	30	49.2	411	4	US-09-252-991A-17658	Sequence 17658, A
150	30	49.2	415	3	US-09-176-657-8	Sequence 8, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and AviHu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0526  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1170 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-313-288B-20

Query Match 91.8%; Score 56; DB 1; Length 1170;  
Best Local Similarity 91.7%; Pred. No. 0.017;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLDVRFVF 12  
| | | | | : | | | |  
Db 208 FQGVLDVRFVF 219

#### RESULT 2

US-09-489-039A-7909  
; Sequence 7909, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 7909  
; LENGTH: 432  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-7909

Query Match 63.9%; Score 39; DB 4; Length 432;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLDVRFVF 12  
| : | : | : | : |  
Db 324 FRGLLEDLRLLF 335

#### RESULT 3

US-08-313-288B-19  
; Sequence 19, Application US/08313288B

SEQ ID NO: 19

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:04:45 ; Search time 12.9515 Seconds  
(without alignments)  
47.833 Million cell updates/sec

Title: US-10-030-735-19  
Perfect score: 61  
Sequence: 1 FQGVQLQDVRFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		Length	DB	ID	Description
	Score	Match				
1	56	91.8	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	39	63.9	432	4	US-09-489-039A-7909	Sequence 7909, Ap
3	38	62.3	1172	1	US-08-313-288B-19	Sequence 19, Appl
4	37	60.7	810	2	US-08-820-170A-34	Sequence 34, Appl
5	37	60.7	810	3	US-09-055-699-34	Sequence 34, Appl
6	37	60.7	810	3	US-09-273-565-34	Sequence 34, Appl
7	37	60.7	810	4	US-09-565-538-34	Sequence 34, Appl
8	37	60.7	810	4	US-09-661-468-34	Sequence 34, Appl
9	37	60.7	810	4	US-09-976-165-34	Sequence 34, Appl
10	36	59.0	162	4	US-09-732-210-527	Sequence 527, App
11	36	59.0	295	3	US-09-150-133-13	Sequence 13, Appl

12	36	59.0	295	3	US-09-150-141-13	Sequence 13, Appl
13	36	59.0	295	3	US-09-374-493-13	Sequence 13, Appl
14	36	59.0	295	3	US-09-374-824-13	Sequence 13, Appl
15	36	59.0	295	3	US-09-374-492-13	Sequence 13, Appl
16	36	59.0	295	4	US-09-785-343-13	Sequence 13, Appl
17	35	57.4	175	4	US-09-252-991A-18826	Sequence 18826, A
18	35	57.4	206	4	US-09-252-991A-20515	Sequence 20515, A
19	35	57.4	421	4	US-08-311-731A-30	Sequence 30, Appl
20	35	57.4	449	4	US-09-489-039A-8009	Sequence 8009, Ap
21	35	57.4	816	2	US-08-820-170A-37	Sequence 37, Appl
22	35	57.4	816	3	US-09-055-699-37	Sequence 37, Appl
23	35	57.4	816	3	US-09-273-565-37	Sequence 37, Appl
24	35	57.4	816	4	US-09-565-538-37	Sequence 37, Appl
25	35	57.4	816	4	US-09-661-468-37	Sequence 37, Appl
26	35	57.4	816	4	US-09-976-165-37	Sequence 37, Appl
27	34	55.7	35	3	US-08-942-686-5	Sequence 5, Appli
28	34	55.7	158	2	US-08-653-402B-6	Sequence 6, Appli
29	34	55.7	169	4	US-09-134-000C-3511	Sequence 3511, Ap
30	34	55.7	273	4	US-09-252-991A-20770	Sequence 20770, A
31	34	55.7	375	4	US-09-134-001C-5050	Sequence 5050, Ap
32	34	55.7	440	1	US-08-307-499-15	Sequence 15, Appl
33	34	55.7	440	3	US-09-299-268-15	Sequence 15, Appl
34	34	55.7	539	4	US-09-800-170-16	Sequence 16, Appl
35	34	55.7	753	3	US-08-942-686-2	Sequence 2, Appli
36	34	55.7	944	4	US-09-134-000C-5578	Sequence 5578, Ap
37	34	55.7	1019	1	US-08-271-364A-7	Sequence 7, Appli
38	34	55.7	1019	2	US-08-222-715B-26	Sequence 26, Appl
39	33	54.1	131	4	US-09-252-991A-22819	Sequence 22819, A
40	33	54.1	156	4	US-09-056-556-227	Sequence 227, App
41	33	54.1	156	4	US-09-072-596-222	Sequence 222, App
42	33	54.1	156	4	US-09-072-967-227	Sequence 227, App
43	33	54.1	187	4	US-09-252-991A-17115	Sequence 17115, A
44	33	54.1	220	4	US-08-134-231C-26	Sequence 26, Appl
45	33	54.1	220	4	US-08-728-160-26	Sequence 26, Appl
46	33	54.1	348	4	US-09-489-039A-13746	Sequence 13746, A
47	33	54.1	408	4	US-09-328-352-6550	Sequence 6550, Ap
48	33	54.1	544	4	US-09-252-991A-30409	Sequence 30409, A
49	33	54.1	596	4	US-09-134-001C-5544	Sequence 5544, Ap
50	33	54.1	659	4	US-09-252-991A-23467	Sequence 23467, A
51	33	54.1	710	4	US-09-252-991A-25700	Sequence 25700, A
52	33	54.1	715	2	US-08-484-993B-10	Sequence 10, Appl
53	33	54.1	715	2	US-08-484-158B-10	Sequence 10, Appl
54	33	54.1	715	2	US-08-484-596A-10	Sequence 10, Appl
55	33	54.1	715	2	US-08-480-150A-10	Sequence 10, Appl
56	33	54.1	715	3	US-08-458-731-10	Sequence 10, Appl
57	33	54.1	715	3	US-08-149-223A-10	Sequence 10, Appl
58	33	54.1	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
59	33	54.1	1242	4	US-09-107-532A-5241	Sequence 5241, Ap
60	33	54.1	1619	4	US-09-392-812A-4	Sequence 4, Appli
61	33	54.1	1704	4	US-09-392-812A-2	Sequence 2, Appli
62	32.5	53.3	563	4	US-09-489-039A-12963	Sequence 12963, A
63	32	52.5	107	4	US-09-198-452A-584	Sequence 584, App
64	32	52.5	132	4	US-09-489-039A-11744	Sequence 11744, A
65	32	52.5	265	4	US-09-252-991A-29034	Sequence 29034, A
66	32	52.5	267	4	US-09-107-532A-4931	Sequence 4931, Ap
67	32	52.5	346	4	US-09-198-452A-60	Sequence 60, Appl
68	32	52.5	358	2	US-08-823-986A-2	Sequence 2, Appli



69	32	52.5	358	3	US-08-784-385-2	Sequence 2, Appli
70	32	52.5	358	5	PCT-US95-10403-2	Sequence 2, Appli
71	32	52.5	368	4	US-09-819-607-5	Sequence 5, Appli
72	32	52.5	390	3	US-08-650-766-7	Sequence 7, Appli
73	32	52.5	390	3	US-08-922-635-6	Sequence 6, Appli
74	32	52.5	390	4	US-09-389-487-7	Sequence 7, Appli
75	32	52.5	420	4	US-09-489-039A-8902	Sequence 8902, Ap
76	32	52.5	497	4	US-09-134-000C-4486	Sequence 4486, Ap
77	32	52.5	550	4	US-09-907-794A-227	Sequence 227, App
78	32	52.5	550	4	US-09-905-125A-227	Sequence 227, App
79	32	52.5	550	4	US-09-902-775A-227	Sequence 227, App
80	32	52.5	559	4	US-09-364-206-47	Sequence 47, Appl
81	32	52.5	596	4	US-09-134-000C-6209	Sequence 6209, Ap
82	32	52.5	651	3	US-08-650-766-6	Sequence 6, Appli
83	32	52.5	651	3	US-08-922-635-5	Sequence 5, Appli
84	32	52.5	651	4	US-09-389-487-6	Sequence 6, Appli
85	32	52.5	667	4	US-09-252-991A-32588	Sequence 32588, A
86	32	52.5	1022	1	US-08-271-364A-8	Sequence 8, Appli
87	32	52.5	1022	2	US-08-222-715B-27	Sequence 27, Appl
88	32	52.5	1070	3	US-08-922-635-22	Sequence 22, Appl
89	32	52.5	1288	3	US-08-762-428A-6	Sequence 6, Appli
90	32	52.5	1504	4	US-09-364-206-2	Sequence 2, Appli
91	31	50.8	129	4	US-09-543-681A-7608	Sequence 7608, Ap
92	31	50.8	229	3	US-09-045-973-7	Sequence 7, Appli
93	31	50.8	242	4	US-09-134-001C-2868	Sequence 2868, Ap
94	31	50.8	249	4	US-09-252-991A-21894	Sequence 21894, A
95	31	50.8	253	4	US-09-489-039A-13477	Sequence 13477, A
96	31	50.8	279	5	PCT-US91-00899-8	Sequence 8, Appli
97	31	50.8	313	4	US-09-540-236-2969	Sequence 2969, Ap
98	31	50.8	369	4	US-09-489-039A-8053	Sequence 8053, Ap
99	31	50.8	371	2	US-08-378-617A-10	Sequence 10, Appl
100	31	50.8	371	2	US-08-378-617A-12	Sequence 12, Appl
101	31	50.8	394	1	US-07-914-281-4	Sequence 4, Appli
102	31	50.8	394	1	US-08-393-246-4	Sequence 4, Appli
103	31	50.8	394	1	US-08-525-058A-4	Sequence 4, Appli
104	31	50.8	394	2	US-08-696-731-4	Sequence 4, Appli
105	31	50.8	394	3	US-09-042-531-4	Sequence 4, Appli
106	31	50.8	394	4	US-08-379-040-2	Sequence 2, Appli
107	31	50.8	394	5	PCT-US91-00899-9	Sequence 9, Appli
108	31	50.8	409	4	US-09-489-039A-14217	Sequence 14217, A
109	31	50.8	420	4	US-09-252-991A-19389	Sequence 19389, A
110	31	50.8	499	4	US-09-107-532A-5524	Sequence 5524, Ap
111	31	50.8	516	4	US-09-291-170A-1	Sequence 1, Appli
112	31	50.8	516	4	US-09-724-884-1	Sequence 1, Appli
113	31	50.8	541	3	US-08-687-590-28	Sequence 28, Appl
114	31	50.8	541	4	US-09-311-311C-25	Sequence 25, Appl
115	31	50.8	605	4	US-09-252-991A-19861	Sequence 19861, A
116	31	50.8	618	4	US-09-134-000C-6538	Sequence 6538, Ap
117	31	50.8	620	3	US-09-232-200-61	Sequence 61, Appl
118	31	50.8	620	4	US-09-232-197-61	Sequence 61, Appl
119	31	50.8	620	4	US-09-232-201-61	Sequence 61, Appl
120	31	50.8	620	4	US-09-232-195-61	Sequence 61, Appl
121	31	50.8	620	4	US-09-418-963-4	Sequence 4, Appli
122	31	50.8	634	4	US-09-252-991A-24131	Sequence 24131, A
123	31	50.8	884	4	US-09-543-681A-5437	Sequence 5437, Ap
124	30	49.2	113	4	US-09-134-001C-5539	Sequence 5539, Ap
125	30	49.2	131	4	US-09-732-210-1647	Sequence 1647, Ap

126	30	49.2	153	2	US-08-387-942C-51	Sequence 51, Appl
127	30	49.2	186	4	US-09-328-352-7876	Sequence 7876, Ap
128	30	49.2	203	4	US-09-252-991A-24219	Sequence 24219, A
129	30	49.2	207	4	US-09-252-991A-31339	Sequence 31339, A
130	30	49.2	218	4	US-08-849-764C-3	Sequence 3, Appli
131	30	49.2	218	4	US-09-262-087-3	Sequence 3, Appli
132	30	49.2	218	4	US-08-463-261B-9	Sequence 9, Appli
133	30	49.2	220	1	US-08-588-163-3	Sequence 3, Appli
134	30	49.2	220	2	US-09-111-070-3	Sequence 3, Appli
135	30	49.2	220	4	US-09-540-530-2	Sequence 2, Appli
136	30	49.2	220	4	US-08-134-231C-27	Sequence 27, Appl
137	30	49.2	220	4	US-09-660-107-11	Sequence 11, Appl
138	30	49.2	220	4	US-08-728-160-27	Sequence 27, Appl
139	30	49.2	226	3	US-09-176-657-3	Sequence 3, Appli
140	30	49.2	226	4	US-09-421-299-3	Sequence 3, Appli
141	30	49.2	265	4	US-09-540-236-3285	Sequence 3285, Ap
142	30	49.2	298	4	US-09-489-039A-8600	Sequence 8600, Ap
143	30	49.2	308	3	US-08-952-089A-13	Sequence 13, Appl
144	30	49.2	330	4	US-09-543-681A-7116	Sequence 7116, Ap
145	30	49.2	341	4	US-09-252-991A-19640	Sequence 19640, A
146	30	49.2	346	4	US-09-134-000C-6034	Sequence 6034, Ap
147	30	49.2	379	4	US-09-491-577-4	Sequence 4, Appli
148	30	49.2	409	4	US-09-328-352-4570	Sequence 4570, Ap
149	30	49.2	411	4	US-09-252-991A-17658	Sequence 17658, A
150	30	49.2	415	3	US-09-176-657-8	Sequence 8, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-313-288B-20

Query Match 91.8%; Score 56; DB 1; Length 1170;  
 Best Local Similarity 91.7%; Pred. No. 0.017;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLDVRFVF 12  
 |||||:||||  
 Db 208 FQGVLDVRFVF 219

#### RESULT 2

US-09-489-039A-7909  
 ; Sequence 7909, Application US/09489039A  
 ; Patent No. 6610836  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary Breton et. al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 KLEBSIELLA  
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 2709.2004001  
 ; CURRENT APPLICATION NUMBER: US/09/489,039A  
 ; CURRENT FILING DATE: 2000-01-27  
 ; PRIOR APPLICATION NUMBER: US 60/117,747  
 ; PRIOR FILING DATE: 1999-01-29  
 ; NUMBER OF SEQ ID NOS: 14342  
 ; SEQ ID NO 7909  
 ; LENGTH: 432  
 ; TYPE: PRT  
 ; ORGANISM: Klebsiella pneumoniae  
 US-09-489-039A-7909

Query Match 63.9%; Score 39; DB 4; Length 432;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLDVRFVF 12  
 |:|:|:|:| :|  
 Db 324 FRGLLEDLRLLF 335

#### RESULT 3

US-08-313-288B-19  
 ; Sequence 19, Application US/08313288B

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:03:50 ; Search time 9.67401 Seconds  
 (without alignments)  
 119.320 Million cell updates/sec

Title: US-10-030-735-19  
 Perfect score: 61  
 Sequence: 1 FQGVLDVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result			%		Query		ID	Description
	No.	Score	Match	Length	DB			
1	56	91.8	229	2	S57957		thrombospondin 1 -	
2	56	91.8	1170	1	TSHUP1		thrombospondin 1 p	
3	56	91.8	1170	2	A40558		thrombospondin 1 p	
4	44	72.1	212	2	D81929		probable imidazole	
5	39	63.9	106	2	C90261		hypothetical prote	
6	38	62.3	298	2	C87403		FdhD protein [impo	
7	38	62.3	326	2	H83832		hypothetical prote	
8	38	62.3	511	2	S77350		hypothetical prote	
9	38	62.3	1172	1	TSHUP2		thrombospondin 2 p	
10	38	62.3	1172	2	A42587		thrombospondin 2 p	
11	37	60.7	145	2	T20985		hypothetical prote	
12	37	60.7	162	2	B88349		protein F15D4.3 [i	
13	37	60.7	810	2	T10756		Nel-homolog protei	

14	36	59.0	163	2	E64083	ribosomal protein
15	36	59.0	248	2	F69113	hypothetical prote
16	36	59.0	295	2	I53296	testis-specific es
17	36	59.0	505	2	E83918	carboxypeptidase B
18	36	59.0	1071	2	T43255	tricorn proteinase
19	36	59.0	1178	1	A39804	thrombospondin pre
20	35	57.4	151	2	C57253	tRNA-pseudouridine
21	35	57.4	155	2	B83124	hypothetical prote
22	35	57.4	386	1	PWWT6	H <sup>+</sup> -transporting tw
23	35	57.4	386	2	JC2100	H <sup>+</sup> -transporting tw
24	35	57.4	421	2	G87199	aspartokinase [imp
25	35	57.4	438	2	C82340	ATP-dependent RNA
26	35	57.4	446	2	D89811	hypothetical prote
27	35	57.4	494	2	S11305	cytochrome P450 2B
28	35	57.4	788	2	T07667	glycogen(starch) s
29	35	57.4	835	2	JP0076	nel protein - chic
30	35	57.4	2139	2	A35672	crumbs protein - f
31	34	55.7	159	2	F95001	transcription regu
32	34	55.7	159	2	H95234	transcription regu
33	34	55.7	159	2	E97873	hypothetical prote
34	34	55.7	159	2	A98099	hypothetical prote
35	34	55.7	195	2	E95110	ABC transporter, A
36	34	55.7	324	2	C87271	general secretion
37	34	55.7	336	2	T38902	probable cinnamoyl
38	34	55.7	336	2	T44795	methyltransferase
39	34	55.7	349	2	C96738	unknown protein F3
40	34	55.7	383	2	T52651	probable topoisome
41	34	55.7	400	2	T24890	hypothetical prote
42	34	55.7	439	2	T23627	hypothetical prote
43	34	55.7	445	2	E86382	hypothetical prote
44	34	55.7	483	2	I40716	H <sup>+</sup> -transporting tw
45	34	55.7	536	2	S75614	hypothetical prote
46	34	55.7	607	1	C71367	probable ATP-depen
47	34	55.7	721	2	T46037	hypothetical prote
48	34	55.7	747	2	AE2929	two component resp
49	34	55.7	783	2	A98353	probable transcrip
50	34	55.7	893	2	T36795	probable penicilli
51	34	55.7	1038	2	T30304	protein RanBP7 - A
52	34	55.7	1312	2	S68593	DNA-directed DNA p
53	34	55.7	1670	2	S71551	DNA-directed DNA p
54	33	54.1	77	2	A82970	hypothetical prote
55	33	54.1	82	2	F96625	hypothetical prote
56	33	54.1	128	2	F83168	probable iron-bind
57	33	54.1	212	2	B81177	amidotransferase H
58	33	54.1	220	1	A35996	metalloproteinase
59	33	54.1	254	2	T20882	hypothetical prote
60	33	54.1	265	2	H64786	ybcM protein - Esc
61	33	54.1	291	2	G86903	cell shape determi
62	33	54.1	306	2	D84416	cytochrome oxidase
63	33	54.1	397	2	E90167	hypothetical prote
64	33	54.1	411	2	G83353	conserved hypothet
65	33	54.1	413	2	S60930	probable membrane
66	33	54.1	417	2	H83708	hypothetical prote
67	33	54.1	421	2	F70794	probable ask prote
68	33	54.1	459	2	B82416	hypothetical prote
69	33	54.1	467	2	D84938	H <sup>+</sup> -transporting tw
70	33	54.1	469	2	T33595	hypothetical prote

71	33	54.1	511	2	AE3405	hypothetical prote
72	33	54.1	586	2	AI2065	hypothetical prote
73	33	54.1	596	2	G82998	probable acyl-CoA
74	33	54.1	603	2	B64444	hypothetical prote
75	33	54.1	715	2	S70397	zona pellucida gly
76	33	54.1	738	2	S01892	hemolysin A precur
77	33	54.1	741	2	A41462	cytolysin/hemolysi
78	33	54.1	741	2	A82486	hemolysin VCA0219
79	33	54.1	943	2	A82934	excinuclease ABC c
80	33	54.1	948	2	H89848	exinuclease ABC su
81	33	54.1	957	2	F69729	excinuclease ABC c
82	33	54.1	957	2	B84099	excinuclease ABC (
83	33	54.1	1142	2	T39103	probable negative
84	33	54.1	3013	2	AB0480	probable invasin Y
85	32	52.5	106	2	F86558	L21 ribosomal prot
86	32	52.5	106	2	C72065	ribosomal protein
87	32	52.5	111	2	S52596	probable membrane
88	32	52.5	126	2	B85085	probable lipid tra
89	32	52.5	128	2	H65029	iron-sulfur cofact
90	32	52.5	128	2	T44282	iron-sulfur cofact
91	32	52.5	128	2	C91053	NifU-like protein
92	32	52.5	128	2	G85897	NifU-like protein
93	32	52.5	128	2	AE0824	NifU-like protein
94	32	52.5	165	1	S69049	iron-sulfur cofact
95	32	52.5	176	2	AG2817	transcription anti
96	32	52.5	176	2	H97595	nusg (AF176664) [i
97	32	52.5	192	2	T50169	iron-sulfur cofact
98	32	52.5	203	2	T25609	hypothetical prote
99	32	52.5	210	2	H81696	hypothetical prote
100	32	52.5	221	2	S50476	hypothetical prote
101	32	52.5	285	2	AC0783	endonuclease IV [i
102	32	52.5	295	2	E72462	probable lipoic ac
103	32	52.5	311	2	F70184	ribose/galactose A
104	32	52.5	320	2	T09288	late embryonic abu
105	32	52.5	358	2	D86289	GA4 protein [impor
106	32	52.5	362	2	T09886	hypothetical prote
107	32	52.5	368	2	T42748	probable threonine
108	32	52.5	379	2	S55900	DNAJ-like protein
109	32	52.5	379	2	T41633	psi protein - fiss
110	32	52.5	382	2	F97013	metal-dependent am
111	32	52.5	385	2	S56224	hypothetical prote
112	32	52.5	385	2	S46532	polygalacturonase
113	32	52.5	393	2	AC1142	N-acyl-L-amino aci
114	32	52.5	396	2	AG2881	HlyD family secret
115	32	52.5	396	2	G97657	hypothetical prote
116	32	52.5	449	2	C86496	hypothetical prote
117	32	52.5	449	2	D72127	hypothetical prote
118	32	52.5	449	2	A81544	hypothetical prote
119	32	52.5	465	2	AC0347	probable membrane
120	32	52.5	467	2	A99494	thermostable carbo
121	32	52.5	471	2	S45068	53K glycoprotein -
122	32	52.5	514	2	T39213	threonine synthase
123	32	52.5	514	2	S49036	threonine synthase
124	32	52.5	515	2	D90108	chaperonin-contain
125	32	52.5	520	2	B96517	hypothetical prote
126	32	52.5	580	2	T21493	hypothetical prote
127	32	52.5	582	2	T07953	lectin-like protei

128	32	52.5	582	2	T07952	lectin-like protei
129	32	52.5	636	2	T10569	probable serine/th
130	32	52.5	661	2	T32518	hypothetical prote
131	32	52.5	728	2	F72693	probable phosphoes
132	32	52.5	830	2	S56940	factor arrest prot
133	32	52.5	968	2	G81743	preprotein translo
134	32	52.5	969	2	G71482	probable protein t
135	32	52.5	993	2	T17230	hypothetical prote
136	32	52.5	1217	2	H89863	hypothetical prote
137	32	52.5	1288	2	T42756	5-oxoprolinase (AT
138	32	52.5	1459	2	G86457	unknown protein, 4
139	32	52.5	1702	2	S42459	DNA-directed DNA p
140	32	52.5	2388	2	JE0271	beta spectrin, bet
141	32	52.5	4936	2	AH2515	hypothetical prote
142	31.5	51.6	144	2	G69425	hypothetical prote
143	31.5	51.6	443	2	AB2719	conserved hypothet
144	31.5	51.6	480	2	F97500	tolerance to colic
145	31	50.8	62	2	D69894	hypothetical prote
146	31	50.8	94	2	AD0805	conserved hypothet
147	31	50.8	102	2	D81910	hypothetical prote
148	31	50.8	108	2	G81370	probable ribonucle
149	31	50.8	120	2	S07727	NADH2 dehydrogenas
150	31	50.8	130	2	E82647	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 91.8%; Score 56; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0017;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 FQGVLDVRFVF 12
          |||||:|||||
Db     190 FQGVLDVRFVF 201

```

##### RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.

A;Reference number: A26155; MUID:87057617; PMID:2430973

A;Accession: A26155

A;Molecule type: mRNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.

A;Reference number: A34274; MUID:89291870; PMID:2544587

A;Accession: A34274

A;Molecule type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.

J. Cell Biol. 108, 729-736, 1989

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>

A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-397 <KOB>

A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276; PMID:3461443

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-374, 'RC' <DIX>

A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992



OM protein - protein search, using sw model

Run on: April 7, 2004, 18:00:05 ; Search time 5.4978 Seconds  
 (without alignments)  
 113.653 Million cell updates/sec

Title: US-10-030-735-19  
 Perfect score: 61  
 Sequence: 1 FQGVLDVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%					
No.	Score	Query	Match	Length	DB	ID	Description
1	56	91.8	1170	1	TSP1_BOVIN	Q28178	bos taurus
2	56	91.8	1170	1	TSP1_HUMAN	P07996	homo sapien
3	56	91.8	1170	1	TSP1_MOUSE	P35441	mus musculu
4	56	91.8	1173	1	TSP1_XENLA	P35448	xenopus lae
5	44	72.1	212	1	HIS5_NEIMA	Q9jvh3	neisseria m
6	38	62.3	1172	1	TSP2_HUMAN	P35442	homo sapien
7	38	62.3	1172	1	TSP2_MOUSE	Q03350	mus musculu
8	37	60.7	810	1	NEL1_HUMAN	Q92832	homo sapien
9	37	60.7	810	1	NEL1_RAT	Q62919	rattus norv
10	37	60.7	894	1	OPLA_HUMAN	O14841	homo sapien
11	36	59.0	162	1	RL10_HAEIN	P44350	haemophilus
12	36	59.0	163	1	RL10_PASMU	Q9ck89	pasteurella
13	36	59.0	295	1	SUOT_MOUSE	P49891	mus musculu
14	36	59.0	1071	1	TRI_THEAC	P96086	thermoplasm
15	36	59.0	1170	1	TSP2_BOVIN	Q95116	bos taurus
16	36	59.0	1178	1	TSP2_CHICK	P35440	gallus gall
17	35	57.4	386	1	ATP6_WHEAT	P20599	triticum ae

18	35	57.4	435	1	RHLB_VIBVU	Q8ddn6	vibrio vuln
19	35	57.4	435	1	RHLB_VIBVY	Q7m9p7	vibrio vuln
20	35	57.4	437	1	RHLB_VIBPA	Q87kh5	vibrio para
21	35	57.4	438	1	RHLB_VIBCH	Q9kv52	vibrio chol
22	35	57.4	439	1	RHLB_SHEON	Q8ejq5	shewanella
23	35	57.4	494	1	CPBB_CANFA	P24460	canis famil
24	35	57.4	788	1	UGS3_SOLTU	Q43847	solanum tub
25	35	57.4	816	1	NEL2_HUMAN	Q99435	homo sapien
26	35	57.4	816	1	NEL2_RAT	Q62918	rattus norv
27	35	57.4	816	1	NEL_CHICK	Q90827	gallus gall
28	35	57.4	910	1	SYL_THEAC	Q9hk31	thermoplasm
29	35	57.4	2139	1	CRB_DROME	P10040	drosophila
30	35	57.4	3579	1	STAN_DROME	Q9v5n8	drosophila
31	34	55.7	211	1	HIS5_CORGL	O69043	corynebacte
32	34	55.7	440	1	KRF1_SPVKA	P32216	swinepox vi
33	34	55.7	483	1	ATPB_CORGL	P42464	corynebacte
34	34	55.7	536	1	YC42_SYNY3	P42349	synechocyst
35	34	55.7	882	1	LY10_HUMAN	Q13342	homo sapien
36	34	55.7	1281	1	TRPS_HUMAN	Q9uhf7	homo sapien
37	34	55.7	1312	1	DPOL_PYRSD	Q51334	pyrococcus
38	34	55.7	1376	1	CRBH_HUMAN	P82279	homo sapien
39	34	55.7	1671	1	DPOL_PYRKO	P77933	pyrococcus
40	34	55.7	1699	1	DPOL_THEG8	Q9hh84	thermococcu
41	33	54.1	194	1	TIM2_RABIT	Q9trz7	oryctolagus
42	33	54.1	212	1	HIS5_NEIMB	Q9k0h2	neisseria m
43	33	54.1	214	1	KGUA_PSESM	Q88be2	pseudomonas
44	33	54.1	217	1	EFTS_PROMA	Q7vcb5	prochloroco
45	33	54.1	220	1	TIM2_BOVIN	P16368	bos taurus
46	33	54.1	220	1	TIM2_CHICK	O42146	gallus gall
47	33	54.1	265	1	YBCM_ECOLI	P77634	escherichia
48	33	54.1	397	1	CD61_SULSO	Q980n4	sulfolobus
49	33	54.1	421	1	AK_MYCTU	P97048	mycobacteri
50	33	54.1	467	1	FLII_BUCAI	P57178	buchnera ap
51	33	54.1	603	1	YB55_METJA	Q58555	methanococc
52	33	54.1	715	1	ZP2_CANFA	P47983	canis famil
53	33	54.1	741	1	HLYA_VIBCH	P09545	vibrio chol
54	33	54.1	816	1	NEL2_MOUSE	Q61220	mus musculu
55	33	54.1	943	1	UVRA_UREPA	Q9pr42	ureaplasma
56	33	54.1	944	1	UVRA_STAEP	Q8cpy9	staphylococ
57	33	54.1	948	1	UVRA_STAAM	Q99vl6	staphylococ
58	33	54.1	948	1	UVRA_STAAW	Q8nxl9	staphylococ
59	33	54.1	957	1	UVRA_BACHD	Q9k6y0	bacillus ha
60	33	54.1	957	1	UVRA_BACSU	O34863	bacillus su
61	33	54.1	958	1	UVRA_OCEIH	Q8enj6	oceanobacil
62	33	54.1	983	1	B3A4_HUMAN	Q96q91	homo sapien
63	32	52.5	106	1	RL21_CHLPN	Q9z806	chlamydia p
64	32	52.5	128	1	NIFU_ECOLI	P77310	escherichia
65	32	52.5	221	1	YEK8_YEAST	P40014	saccharomyc
66	32	52.5	285	1	END4_SALTI	Q8z593	salmonella
67	32	52.5	295	1	LIPA_AERPE	Q9y9e3	aeropyrum p
68	32	52.5	338	1	ETFA_MEGEL	O85692	megasphaera
69	32	52.5	379	1	PSI_SCHPO	Q09912	schizosacch
70	32	52.5	385	1	YFD0_YEAST	P43567	saccharomyc
71	32	52.5	461	1	ARLY_NOSPU	Q9lae5	nostoc punc
72	32	52.5	514	1	THRC_SCHPO	Q42598	schizosacch
73	32	52.5	550	1	COCH_HUMAN	O43405	homo sapien
74	32	52.5	552	1	COCH_MOUSE	Q62507	mus musculu

75	32	52.5	830	1	FAR1_YEAST	P21268	saccharomyc
76	32	52.5	953	1	B3A4_RAT	Q8k4v2	rattus norv
77	32	52.5	955	1	B3A4_RABIT	Q9gky1	oryctolagus
78	32	52.5	968	1	SECA_CHLMU	Q9plm5	chlamydia m
79	32	52.5	969	1	SECA_CHLTR	O84707	chlamydia t
80	32	52.5	1288	1	OPLA_MOUSE	Q8k010	mus musculu
81	32	52.5	1288	1	OPLA_RAT	P97608	rattus norv
82	32	52.5	1702	1	DPOL_THELI	P30317	thermococcu
83	32	52.5	2388	1	SPCP_RAT	Q9qwn8	rattus norv
84	32	52.5	2390	1	SPCP_HUMAN	O15020	homo sapien
85	31.5	51.6	144	1	YE08_ARCFU	O28864	archaeoglob
86	31	50.8	62	1	YNZB_BACSU	O31807	bacillus su
87	31	50.8	108	1	RNPA_CAMJE	Q9pnx5	campylobact
88	31	50.8	120	1	NU3M_PARTE	P15579	paramecium
89	31	50.8	146	1	MRAZ_RHIME	Q92nl3	rhizobium m
90	31	50.8	181	1	CSL4_SCHPO	O59821	schizosacch
91	31	50.8	196	1	PHOU_ARCFU	O28911	archaeoglob
92	31	50.8	199	1	HIS5_HAEIN	P44340	haemophilus
93	31	50.8	200	1	REP2_ZYGFE	P13742	zygosacchar
94	31	50.8	202	1	KGUA_SCHPO	Q9p6i5	schizosacch
95	31	50.8	230	1	KCY_MYCTU	O33211	mycobacteri
96	31	50.8	248	1	CUTC_ECOLI	P46719	escherichia
97	31	50.8	313	1	RLA0_NEUCR	Q96tj5	neurospora
98	31	50.8	371	1	GATR_PIG	P50127	sus scrofa
99	31	50.8	394	1	GATR_MOUSE	P23336	mus musculu
100	31	50.8	404	1	ISCS_HAEIN	Q57337	haemophilus
101	31	50.8	405	1	ARGJ_PSEAE	Q9hw04	pseudomonas
102	31	50.8	405	1	KRF1_VACCP	P29884	vaccinia vi
103	31	50.8	407	1	AATC_ORYSA	P37833	oryza sativ
104	31	50.8	415	1	V034_FOWPV	Q9j5g9	fowlpox vir
105	31	50.8	417	1	PROA_HAEIN	P45121	haemophilus
106	31	50.8	418	1	AAT1_MEDSA	P28011	medicago sa
107	31	50.8	418	1	ARP6_CAEEL	Q09443	caenorhabdi
108	31	50.8	421	1	AK_MYCSM	P41403	mycobacteri
109	31	50.8	439	1	KRF1_VACCC	P21095	vaccinia vi
110	31	50.8	439	1	KRF1_VARV	P33801	variola vir
111	31	50.8	449	1	AAT3_ARATH	P46644	arabidopsis
112	31	50.8	501	1	YPWA_BACSU	P50848	bacillus su
113	31	50.8	541	1	TCPE_MOUSE	P80316	mus musculu
114	31	50.8	543	1	RHLB_XYLFA	Q9pa24	xylella fas
115	31	50.8	543	1	RHLB_XYLFT	Q879y6	xylella fas
116	31	50.8	550	1	RT03_OENBE	P27754	oenothera b
117	31	50.8	576	1	YN15_YEAST	P53838	saccharomyc
118	31	50.8	577	1	ALG8_YEAST	P40351	saccharomyc
119	31	50.8	603	1	PLK1_RAT	Q62673	rattus norv
120	31	50.8	610	1	U171_DROME	Q9vub4	drosophila
121	31	50.8	616	1	RFX5_HUMAN	P48382	homo sapien
122	31	50.8	620	1	VLCS_RAT	P97524	rattus norv
123	31	50.8	672	1	HYFB_ECOLI	P23482	escherichia
124	31	50.8	738	1	S521_RAT	Q9qy02	rattus norv
125	31	50.8	753	1	ZIN_HUMAN	Q9nrl3	homo sapien
126	31	50.8	817	1	KPRO_MAIZE	P17801	zea mays (m
127	31	50.8	819	1	YCXA_CHLVU	P56370	chlorella v
128	31	50.8	835	1	SYFB_MYCLE	Q9cc16	mycobacteri
129	31	50.8	899	1	VP3_EHDV1	P27281	epizootic h
130	31	50.8	935	1	UVRA_HELPY	P56474	helicobacte
131	31	50.8	937	1	DBS_RAT	Q63406	rattus norv

132	31	50.8	941	1	UVRA_HELPJ	Q9zld6 helicobacte
133	31	50.8	960	1	UVRA_TREPA	O83527 treponema p
134	31	50.8	1108	1	DBS_HUMAN	O15068 homo sapien
135	31	50.8	1149	1	DBS_MOUSE	Q64096 mus musculu
136	31	50.8	1177	1	SP97_DICDI	Q95zg3 dictyosteli
137	31	50.8	1177	1	Y307_MYCGE	P47549 mycoplasma
138	31	50.8	1829	1	DPOL_THEST	O33845 thermococcu
139	31	50.8	1835	1	DUR1_YEAST	P32528 saccharomyc
140	31	50.8	1995	1	YCX7_CHLRE	P36495 chlamydomon
141	31	50.8	2210	1	RRPO_TACV	P20430 tacaribe vi
142	30.5	50.0	397	1	Y303_ARCFU	O29939 archaeoglob
143	30	49.2	79	1	DCOR_PARBR	Q92445 paracoccidi
144	30	49.2	94	1	NI8M_NEUCR	Q07842 neurospora
145	30	49.2	131	1	RS8_BACSU	P12879 bacillus su
146	30	49.2	206	1	HIS5_MYCBO	P59957 mycobacteri
147	30	49.2	206	1	HIS5_MYCTU	O06589 mycobacteri
148	30	49.2	217	1	RIT2_MOUSE	P70425 mus musculu
149	30	49.2	220	1	TIM2_CAVPO	Q9wuc6 cavia porce
150	30	49.2	220	1	TIM2_HUMAN	P16035 homo sapien

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.  
AC Q28178; Q28179;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Thrombospondin 1 precursor.  
GN THBS1 OR TSP1 OR TSP-1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Holstein; TISSUE=Tooth;  
RX MEDLINE=98173773; PubMed=9507054;  
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,  
RA Inoue H.;  
RT "cdNA cloning of bovine thrombospondin 1 and its expression in  
RT odontoblasts and predentin.";  
RL Biochim. Biophys. Acta 1382:17-22(1998).  
RN [2]  
RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.  
RC TISSUE=Aortic endothelium;  
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-  
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis  
CC and/or maintenance of dentin and dental pulp.

```

CC  -!- SUBUNIT: Homotrimer; disulfide-linked.
CC  -!- TISSUE SPECIFICITY: Odontoblasts.
CC  -!- SIMILARITY: Belongs to the thrombospondin family.
CC  -!- SIMILARITY: Contains 1 VWFC domain.
CC  -!- SIMILARITY: Contains 3 EGF-like domains.
CC  -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC  -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC  -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AB005287; BAA21115.1; -.
DR  EMBL; X87618; CAA60950.1; -.
DR  EMBL; X87619; CAA60951.1; -.
DR  PIR; S55501; S55501.
DR  GlycoSuiteDB; Q28178; -.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR006209; EGF_like.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000884; TSP1.
DR  InterPro; IPR008085; TSP_1.
DR  InterPro; IPR003367; tsp_3.
DR  InterPro; IPR008859; TSPC.
DR  InterPro; IPR003129; TSPN.
DR  InterPro; IPR001007; VWF_C.
DR  Pfam; PF00008; EGF; 2.
DR  Pfam; PF00090; tsp_1; 3.
DR  Pfam; PF02412; tsp_3; 13.
DR  Pfam; PF05735; TSPC; 1.
DR  Pfam; PF02210; TSPN; 1.
DR  Pfam; PF00093; vwc; 1.
DR  PRINTS; PR01705; TSP1REPEAT.
DR  SMART; SM00181; EGF; 3.
DR  SMART; SM00209; TSP1; 3.
DR  SMART; SM00210; TSPN; 1.
DR  SMART; SM00214; VWC; 1.
DR  PROSITE; PS00022; EGF_1; FALSE_NEG.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS50026; EGF_3; 2.
DR  PROSITE; PS50092; TSP1; 3.
DR  PROSITE; PS01208; VWFC_1; 1.
DR  PROSITE; PS50184; VWFC_2; 1.
KW  Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW  EGF-like domain; Signal.
FT  SIGNAL          1      18      BY SIMILARITY.
FT  CHAIN           19     1170     THROMBOSPONDIN 1.
FT  DOMAIN          19     232     HEPARIN-BINDING (POTENTIAL).
FT  DOMAIN          24     221     TSP N-TERMINAL.
FT  DOMAIN          316    373     VWFC.
FT  DOMAIN          379    429     TSP TYPE-1 1.
FT  DOMAIN          435    490     TSP TYPE-1 2.

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:00:35 ; Search time 31.5595 Seconds  
 (without alignments)  
 119.971 Million cell updates/sec

Title: US-10-030-735-19  
 Perfect score: 61  
 Sequence: 1 FQGVLDVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertibrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Query	Match Length DB	ID	Description
No.					

-----

1	56	91.8	229	6	Q28194	Q28194 bos taurus
2	56	91.8	496	13	Q7SY84	Q7sy84 xenopus lae
3	56	91.8	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	56	91.8	1171	11	Q80YQ1	Q80yq1 mus musculu
5	39	63.9	106	17	Q97Z50	Q97z50 sulfolobus
6	39	63.9	134	9	Q9G0G7	Q9g0g7 roseophage
7	38	62.3	298	16	Q9A8V9	Q9a8v9 caulobacter
8	38	62.3	326	16	Q9KCV5	Q9kcv5 bacillus ha
9	38	62.3	346	16	Q8FFZ0	Q8ffz0 escherichia
10	38	62.3	389	11	Q8BGR4	Q8bgr4 mus musculu
11	38	62.3	511	16	P73413	P73413 synechocyst
12	38	62.3	844	11	Q7TSG9	Q7tsg9 mus musculu
13	38	62.3	1172	11	Q8CG21	Q8cg21 mus musculu
14	38	62.3	1172	11	Q7TMT3	Q7tmt3 mus musculu
15	38	62.3	1405	11	Q8VHS2	Q8vhs2 mus musculu
16	38	62.3	7716	16	Q7UWZ8	Q7uwz8 rhodopirell
17	37	60.7	135	2	Q8VNW4	Q8vnw4 comamonas a
18	37	60.7	145	5	Q93511	Q93511 caenorhabdi
19	37	60.7	146	17	Q976Q1	Q976q1 sulfolobus
20	37	60.7	242	11	Q8BM21	Q8bm21 mus musculu
21	37	60.7	457	16	Q89H49	Q89h49 bradyrhizob
22	37	60.7	517	5	Q8MRJ7	Q8mrj7 drosophila
23	37	60.7	537	10	Q943D6	Q943d6 oryza sativ
24	37	60.7	585	16	Q87IV8	Q87iv8 vibrio para
25	37	60.7	727	17	Q8PRY3	Q8pry3 methanosarc
26	37	60.7	733	17	Q8TLX6	Q8tlx6 methanosarc
27	37	60.7	1039	5	Q9VQB1	Q9vqb1 drosophila
28	36	59.0	86	16	Q7UYA0	Q7uya0 rhodopirell
29	36	59.0	163	16	Q892M4	Q892m4 clostridium
30	36	59.0	248	17	O27872	O27872 methanobact
31	36	59.0	295	11	Q8JZX7	Q8jzx7 mus musculu
32	36	59.0	295	11	Q9D566	Q9d566 mus musculu
33	36	59.0	353	10	Q9FSY1	Q9fsy1 chara fragi
34	36	59.0	365	10	Q9FSY6	Q9fsy6 conocephalu
35	36	59.0	366	10	Q9FSY2	Q9fsy2 chara fragi
36	36	59.0	459	16	Q9CMS7	Q9cms7 pasteurella
37	36	59.0	505	16	Q9KAY7	Q9kay7 bacillus ha
38	36	59.0	637	5	Q9VGE7	Q9vge7 drosophila
39	36	59.0	648	16	Q87TY2	Q87ty2 pseudomonas
40	36	59.0	750	11	Q8BVV4	Q8bvv4 mus musculu
41	35	57.4	155	16	Q9HWJ8	Q9hwj8 pseudomonas
42	35	57.4	210	10	Q9LKS1	Q9lks1 arabidopsis
43	35	57.4	230	10	Q9LI90	Q9li90 arabidopsis
44	35	57.4	247	16	Q81EN4	Q81en4 bacillus ce
45	35	57.4	397	16	Q7VM72	Q7vm72 haemophilus
46	35	57.4	421	16	Q9CB77	Q9cb77 mycobacteri
47	35	57.4	446	16	Q99WE4	Q99we4 staphylococ
48	35	57.4	446	16	Q8NY23	Q8ny23 staphylococ
49	35	57.4	505	16	Q81SQ9	Q81sq9 bacillus an
50	35	57.4	505	16	Q81FL5	Q81fl5 bacillus ce
51	35	57.4	553	6	Q8HXE2	Q8hxe2 macaca fasc
52	35	57.4	683	10	Q9SSG5	Q9ssg5 arabidopsis
53	35	57.4	696	2	Q9F661	Q9f661 burkholderi
54	35	57.4	766	2	Q8KX20	Q8kx20 synechococc
55	35	57.4	814	13	Q7ZXL5	Q7zxl5 xenopus lae
56	35	57.4	815	4	Q96JS2	Q96js2 homo sapien
57	35	57.4	816	11	Q8R417	Q8r417 rattus norv

58	35	57.4	944	3	Q96WT2	Q96wt2 phoma betae
59	35	57.4	1711	9	Q8LTK2	Q8ltk2 lactococcus
60	35	57.4	1713	9	Q94MA1	Q94ma1 lactococcus
61	35	57.4	2146	5	Q9VC97	Q9vc97 drosophila
62	35	57.4	2425	6	Q95MJ0	Q95mj0 tupaia glis
63	35	57.4	4254	16	Q7UU38	Q7uu38 rhodopirell
64	34	55.7	44	2	Q84DX4	Q84dx4 tropheryma
65	34	55.7	133	16	Q7VLK3	Q7vlk3 haemophilus
66	34	55.7	142	16	Q83G93	Q83g93 tropheryma
67	34	55.7	159	2	Q9R2W8	Q9r2w8 streptococc
68	34	55.7	159	16	Q97CV2	Q97cv2 streptococc
69	34	55.7	159	16	Q8CM18	Q8cm18 streptococc
70	34	55.7	195	16	Q97R72	Q97r72 streptococc
71	34	55.7	214	16	Q82ZE3	Q82ze3 enterococcu
72	34	55.7	223	10	Q9LMB5	Q9lmb5 arabidopsis
73	34	55.7	223	10	Q8W4C9	Q8w4c9 arabidopsis
74	34	55.7	259	17	Q8TPP9	Q8tp9 methanosarc
75	34	55.7	264	17	Q8TPQ0	Q8tpq0 methanosarc
76	34	55.7	264	17	Q8PSF7	Q8psf7 methanosarc
77	34	55.7	278	10	Q94BI7	Q94bi7 celosia sp.
78	34	55.7	318	10	Q8LFJ4	Q8lfj4 arabidopsis
79	34	55.7	318	10	Q9FKM1	Q9fkm1 arabidopsis
80	34	55.7	324	16	Q9ABP6	Q9abp6 caulobacter
81	34	55.7	334	16	Q8Y393	Q8y393 ralstonia s
82	34	55.7	336	3	Q9UT59	Q9ut59 schizosacch
83	34	55.7	336	11	Q9QZT1	Q9qzt1 mus musculu
84	34	55.7	345	10	Q94CZ5	Q94cz5 oryza sativ
85	34	55.7	349	10	Q9FVW0	Q9fvw0 arabidopsis
86	34	55.7	361	17	Q974W0	Q974w0 sulfolobus
87	34	55.7	374	16	Q8CNB7	Q8cnb7 staphylococ
88	34	55.7	376	11	Q91WU5	Q91wu5 mus musculu
89	34	55.7	383	10	Q9M4A1	Q9m4a1 arabidopsis
90	34	55.7	390	16	Q7VIB4	Q7vib4 helicobacte
91	34	55.7	394	16	Q87M68	Q87m68 vibrio para
92	34	55.7	400	5	O18086	O18086 caenorhabdi
93	34	55.7	408	12	Q9DL64	Q9dl64 influenza a
94	34	55.7	427	11	Q8BKD8	Q8bkd8 mus musculu
95	34	55.7	439	5	O45689	O45689 caenorhabdi
96	34	55.7	445	3	Q9UVG2	Q9uv92 phaeosphaer
97	34	55.7	445	10	Q9FRH7	Q9frh7 arabidopsis
98	34	55.7	460	10	Q94DF6	Q94df6 oryza sativ
99	34	55.7	481	16	Q8FQ20	Q8fq20 corynebacte
100	34	55.7	490	16	Q7VGZ8	Q7vgz8 helicobacte
101	34	55.7	593	2	Q8GH66	Q8gh66 mycobacteri
102	34	55.7	595	5	Q86NR6	Q86nr6 drosophila
103	34	55.7	607	16	O83141	O83141 treponema p
104	34	55.7	659	16	Q8RCI0	Q8rci0 thermoanaer
105	34	55.7	689	16	Q8R808	Q8r808 thermoanaer
106	34	55.7	700	2	O24781	O24781 bacillus sp
107	34	55.7	721	5	Q9VTH0	Q9vth0 drosophila
108	34	55.7	721	10	Q9M2Y7	Q9m2y7 arabidopsis
109	34	55.7	783	16	Q8UBI1	Q8ubi1 agrobacteri
110	34	55.7	795	16	Q88KD6	Q88kd6 pseudomonas
111	34	55.7	893	16	Q9S258	Q9s258 streptomyce
112	34	55.7	895	11	Q8BYI0	Q8byi0 mus musculu
113	34	55.7	932	11	Q7TN09	Q7tn09 mus musculu
114	34	55.7	956	13	Q8QG47	Q8qg47 aegithalos



115	34	55.7	1010	4	Q9H177	Q9h177 homo sapien
116	34	55.7	1034	5	Q7Z2B9	Q7z2b9 trypanosoma
117	34	55.7	1037	13	Q7ZWZ8	Q7zwz8 xenopus lae
118	34	55.7	1038	4	O95373	O95373 homo sapien
119	34	55.7	1038	11	Q7TQ63	Q7tq63 mus musculu
120	34	55.7	1038	13	O42480	O42480 xenopus lae
121	34	55.7	1039	11	Q9EPL8	Q9epl8 mus musculu
122	34	55.7	1045	5	Q9NKY1	Q9nky1 drosophila
123	34	55.7	1093	3	Q872P9	Q872p9 neurospora
124	34	55.7	1193	5	Q8IMU8	Q8imu8 drosophila
125	34	55.7	1249	16	Q88U41	Q88u41 lactobacill
126	34	55.7	1264	16	Q836J8	Q836j8 enterococcu
127	34	55.7	1406	4	Q8WWY0	Q8wwy0 homo sapien
128	34	55.7	1464	3	Q9C229	Q9c229 neurospora
129	34	55.7	2060	5	Q7YYY7	Q7yyy7 cryptospori
130	34	55.7	4971	5	Q8IBG1	Q8ibg1 plasmodium
131	34	55.7	6307	4	Q8WVG9	Q8wxg9 homo sapien
132	33	54.1	59	10	Q84SP4	Q84sp4 oryza sativ
133	33	54.1	77	16	Q9HTF9	Q9htf9 pseudomonas
134	33	54.1	82	10	Q9ZUI2	Q9zui2 arabidopsis
135	33	54.1	128	16	Q9HXI9	Q9hxi9 pseudomonas
136	33	54.1	128	16	Q88PK7	Q88pk7 pseudomonas
137	33	54.1	133	16	Q8Y0M3	Q8y0m3 ralstonia s
138	33	54.1	136	4	Q15153	Q15153 homo sapien
139	33	54.1	142	2	Q8RLE3	Q8rle3 mycoplasma
140	33	54.1	170	16	Q8EU96	Q8eu96 mycoplasma
141	33	54.1	200	6	Q28465	Q28465 macropus eu
142	33	54.1	213	16	Q8EEA1	Q8eea1 shewanella
143	33	54.1	217	16	Q8RC20	Q8rc20 thermoanaer
144	33	54.1	217	16	Q7VCB5	Q7vcb5 prochloroco
145	33	54.1	218	16	Q7TV13	Q7tv13 prochloroco
146	33	54.1	245	16	Q83R37	Q83r37 shigella fl
147	33	54.1	250	2	Q8GPX5	Q8gpx5 pseudomonas
148	33	54.1	254	5	Q19450	Q19450 caenorhabdi
149	33	54.1	265	16	Q8FGL4	Q8fgl4 escherichia
150	33	54.1	279	10	Q9ZTZ5	Q9ztz5 amaranthus

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.

AC Q28194;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Thrombospondin-1 (Fragment).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lect\_gl.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 91.8%; Score 56; DB 6; Length 229;  
 Best Local Similarity 91.7%; Pred. No. 0.0055;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLDVRFVF 12  
 |||||:|||||  
 Db 190 FQGVLDVRFVF 201

## RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:04:45 ; Search time 12.9515 Seconds  
 (without alignments)  
 47.833 Million cell updates/sec

Title: US-10-030-735-20  
 Perfect score: 60  
 Sequence: 1 FQGVLLQVRVFV 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	55	91.7	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	39	65.0	169	4	US-09-134-000C-3511	Sequence 3511, Ap
3	37	61.7	1172	1	US-08-313-288B-19	Sequence 19, Appl
4	36	60.0	175	4	US-09-252-991A-18826	Sequence 18826, A
5	36	60.0	440	1	US-08-307-499-15	Sequence 15, Appl
6	36	60.0	440	3	US-09-299-268-15	Sequence 15, Appl
7	36	60.0	516	4	US-09-291-170A-1	Sequence 1, Appli
8	36	60.0	516	4	US-09-724-884-1	Sequence 1, Appli
9	35	58.3	1142	4	US-09-106-075A-89	Sequence 89, Appl
10	35	58.3	1333	4	US-09-347-878-20	Sequence 20, Appl
11	34	56.7	539	4	US-09-800-170-16	Sequence 16, Appl

12	34	56.7	710	4	US-09-489-039A-14121	Sequence 14121, A
13	34	56.7	944	4	US-09-134-000C-5578	Sequence 5578, Ap
14	34	56.7	1019	1	US-08-271-364A-7	Sequence 7, Appli
15	34	56.7	1019	2	US-08-222-715B-26	Sequence 26, Appl
16	33	55.0	269	2	US-09-006-535-6	Sequence 6, Appli
17	33	55.0	296	2	US-09-006-535-5	Sequence 5, Appli
18	33	55.0	344	4	US-09-328-352-6058	Sequence 6058, Ap
19	33	55.0	396	4	US-09-252-991A-28379	Sequence 28379, A
20	33	55.0	432	4	US-09-489-039A-7909	Sequence 7909, Ap
21	33	55.0	510	4	US-09-252-991A-25258	Sequence 25258, A
22	33	55.0	715	2	US-08-484-993B-10	Sequence 10, Appl
23	33	55.0	715	2	US-08-484-158B-10	Sequence 10, Appl
24	33	55.0	715	2	US-08-484-596A-10	Sequence 10, Appl
25	33	55.0	715	2	US-08-480-150A-10	Sequence 10, Appl
26	33	55.0	715	3	US-08-458-731-10	Sequence 10, Appl
27	33	55.0	715	3	US-08-149-223A-10	Sequence 10, Appl
28	33	55.0	775	4	US-09-252-991A-22591	Sequence 22591, A
29	33	55.0	816	4	US-09-266-225D-12	Sequence 12, Appl
30	33	55.0	903	4	US-09-252-991A-28775	Sequence 28775, A
31	33	55.0	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
32	33	55.0	1148	4	US-09-106-075A-90	Sequence 90, Appl
33	33	55.0	1242	4	US-09-107-532A-5241	Sequence 5241, Ap
34	32	53.3	177	4	US-09-328-352-5870	Sequence 5870, Ap
35	32	53.3	202	4	US-09-543-681A-7852	Sequence 7852, Ap
36	32	53.3	226	3	US-09-176-657-3	Sequence 3, Appli
37	32	53.3	226	4	US-09-421-299-3	Sequence 3, Appli
38	32	53.3	298	4	US-09-489-039A-12184	Sequence 12184, A
39	32	53.3	338	4	US-09-540-236-3196	Sequence 3196, Ap
40	32	53.3	368	4	US-09-819-607-5	Sequence 5, Appli
41	32	53.3	390	3	US-08-650-766-7	Sequence 7, Appli
42	32	53.3	390	3	US-08-922-635-6	Sequence 6, Appli
43	32	53.3	390	4	US-09-389-487-7	Sequence 7, Appli
44	32	53.3	445	4	US-09-252-991A-18593	Sequence 18593, A
45	32	53.3	559	4	US-09-364-206-47	Sequence 47, Appl
46	32	53.3	624	4	US-09-107-532A-5745	Sequence 5745, Ap
47	32	53.3	634	4	US-09-134-001C-3390	Sequence 3390, Ap
48	32	53.3	651	3	US-08-650-766-6	Sequence 6, Appli
49	32	53.3	651	3	US-08-922-635-5	Sequence 5, Appli
50	32	53.3	651	4	US-09-389-487-6	Sequence 6, Appli
51	32	53.3	906	1	US-08-254-573-2	Sequence 2, Appli
52	32	53.3	906	1	US-08-687-379-2	Sequence 2, Appli
53	32	53.3	906	1	US-08-687-379-4	Sequence 4, Appli
54	32	53.3	906	4	US-08-172-332-1	Sequence 1, Appli
55	32	53.3	906	4	US-08-216-326-2	Sequence 2, Appli
56	32	53.3	907	1	US-07-718-575-2	Sequence 2, Appli
57	32	53.3	907	1	US-08-481-206-2	Sequence 2, Appli
58	32	53.3	907	2	US-08-486-269A-2	Sequence 2, Appli
59	32	53.3	1022	1	US-08-271-364A-8	Sequence 8, Appli
60	32	53.3	1022	2	US-08-222-715B-27	Sequence 27, Appl
61	32	53.3	1070	3	US-08-922-635-22	Sequence 22, Appl
62	32	53.3	1128	4	US-09-252-991A-31032	Sequence 31032, A
63	32	53.3	1207	4	US-09-064-199-16	Sequence 16, Appl
64	32	53.3	1291	4	US-09-252-991A-19504	Sequence 19504, A
65	32	53.3	1504	4	US-09-364-206-2	Sequence 2, Appli
66	32	53.3	15281	2	US-08-471-119A-2	Sequence 2, Appli
67	31.5	52.5	344	4	US-09-252-991A-30083	Sequence 30083, A
68	31	51.7	153	2	US-08-387-942C-51	Sequence 51, Appl

69	31	51.7	286	4	US-09-252-991A-30552	Sequence 30552, A
70	31	51.7	297	2	US-09-006-535-4	Sequence 4, Appli
71	31	51.7	298	2	US-09-006-535-3	Sequence 3, Appli
72	31	51.7	298	4	US-09-134-000C-4916	Sequence 4916, Ap
73	31	51.7	299	4	US-09-648-004-22	Sequence 22, Appl
74	31	51.7	322	4	US-09-543-681A-7868	Sequence 7868, Ap
75	31	51.7	341	4	US-09-593-359-2	Sequence 2, Appli
76	31	51.7	345	4	US-09-134-000C-4688	Sequence 4688, Ap
77	31	51.7	408	4	US-09-252-991A-30312	Sequence 30312, A
78	31	51.7	409	4	US-09-489-039A-14217	Sequence 14217, A
79	31	51.7	419	4	US-09-328-352-4398	Sequence 4398, Ap
80	31	51.7	420	4	US-09-252-991A-19389	Sequence 19389, A
81	31	51.7	445	4	US-09-107-532A-5506	Sequence 5506, Ap
82	31	51.7	503	4	US-09-593-359-4	Sequence 4, Appli
83	31	51.7	541	3	US-08-687-590-28	Sequence 28, Appl
84	31	51.7	541	4	US-09-311-311C-25	Sequence 25, Appl
85	31	51.7	557	1	US-08-424-788-6	Sequence 6, Appli
86	31	51.7	578	1	US-08-424-788-5	Sequence 5, Appli
87	31	51.7	578	1	US-08-110-683-2	Sequence 2, Appli
88	31	51.7	578	2	US-08-683-743-2	Sequence 2, Appli
89	31	51.7	578	2	US-08-477-166-2	Sequence 2, Appli
90	31	51.7	578	2	US-08-472-097-2	Sequence 2, Appli
91	31	51.7	578	4	US-09-439-672-2	Sequence 2, Appli
92	31	51.7	578	5	PCT-US93-11638-2	Sequence 2, Appli
93	31	51.7	591	4	US-09-711-164-425	Sequence 425, App
94	31	51.7	610	4	US-09-489-039A-12251	Sequence 12251, A
95	31	51.7	623	4	US-09-489-039A-11811	Sequence 11811, A
96	31	51.7	728	4	US-09-543-681A-8132	Sequence 8132, Ap
97	31	51.7	765	4	US-09-323-872A-24	Sequence 24, Appl
98	31	51.7	765	4	US-09-323-872A-30	Sequence 30, Appl
99	31	51.7	765	4	US-09-072-433-29	Sequence 29, Appl
100	31	51.7	765	4	US-09-072-433-34	Sequence 34, Appl
101	31	51.7	810	2	US-08-820-170A-34	Sequence 34, Appl
102	31	51.7	810	3	US-09-055-699-34	Sequence 34, Appl
103	31	51.7	810	3	US-09-273-565-34	Sequence 34, Appl
104	31	51.7	810	4	US-09-565-538-34	Sequence 34, Appl
105	31	51.7	810	4	US-09-661-468-34	Sequence 34, Appl
106	31	51.7	810	4	US-09-976-165-34	Sequence 34, Appl
107	31	51.7	825	4	US-09-252-991A-17163	Sequence 17163, A
108	31	51.7	872	2	US-08-387-942C-5	Sequence 5, Appli
109	31	51.7	1049	4	US-09-252-991A-19919	Sequence 19919, A
110	31	51.7	1288	3	US-08-762-428A-6	Sequence 6, Appli
111	31	51.7	1527	4	US-09-376-330-2	Sequence 2, Appli
112	31	51.7	1843	3	US-09-413-814-50	Sequence 50, Appl
113	30	50.0	47	4	US-09-149-476-562	Sequence 562, App
114	30	50.0	74	4	US-09-149-476-408	Sequence 408, App
115	30	50.0	82	4	US-09-198-452A-1025	Sequence 1025, Ap
116	30	50.0	113	4	US-09-134-001C-5539	Sequence 5539, Ap
117	30	50.0	135	4	US-09-328-352-5042	Sequence 5042, Ap
118	30	50.0	153	2	US-08-387-942C-43	Sequence 43, Appl
119	30	50.0	154	4	US-09-732-210-871	Sequence 871, App
120	30	50.0	162	4	US-09-732-210-527	Sequence 527, App
121	30	50.0	186	4	US-09-328-352-7876	Sequence 7876, Ap
122	30	50.0	199	3	US-09-232-191-19	Sequence 19, Appl
123	30	50.0	199	3	US-09-232-200-19	Sequence 19, Appl
124	30	50.0	199	4	US-09-232-197-19	Sequence 19, Appl
125	30	50.0	199	4	US-09-232-201-19	Sequence 19, Appl

126	30	50.0	199	4	US-09-232-195-19	Sequence 19, Appl
127	30	50.0	260	4	US-09-252-991A-25742	Sequence 25742, A
128	30	50.0	295	3	US-09-150-133-13	Sequence 13, Appl
129	30	50.0	295	3	US-09-150-141-13	Sequence 13, Appl
130	30	50.0	295	3	US-09-374-493-13	Sequence 13, Appl
131	30	50.0	295	3	US-09-374-824-13	Sequence 13, Appl
132	30	50.0	295	3	US-09-374-492-13	Sequence 13, Appl
133	30	50.0	295	4	US-09-785-343-13	Sequence 13, Appl
134	30	50.0	308	4	US-09-252-991A-30893	Sequence 30893, A
135	30	50.0	335	4	US-09-252-991A-24116	Sequence 24116, A
136	30	50.0	342	4	US-09-252-991A-16724	Sequence 16724, A
137	30	50.0	354	3	US-09-232-200-55	Sequence 55, Appl
138	30	50.0	354	4	US-09-232-197-55	Sequence 55, Appl
139	30	50.0	354	4	US-09-232-201-55	Sequence 55, Appl
140	30	50.0	354	4	US-09-232-195-55	Sequence 55, Appl
141	30	50.0	365	2	US-08-724-974A-3	Sequence 3, Appli
142	30	50.0	365	4	US-09-170-496D-118	Sequence 118, App
143	30	50.0	365	4	US-09-170-496D-226	Sequence 226, App
144	30	50.0	365	4	US-09-364-425B-27	Sequence 27, Appl
145	30	50.0	420	4	US-08-745-995A-14	Sequence 14, Appl
146	30	50.0	420	4	US-09-005-352-14	Sequence 14, Appl
147	30	50.0	427	4	US-09-550-645-2	Sequence 2, Appli
148	30	50.0	429	4	US-09-219-865B-12	Sequence 12, Appl
149	30	50.0	430	4	US-09-134-001C-4712	Sequence 4712, Ap
150	30	50.0	436	4	US-09-252-991A-28888	Sequence 28888, A

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

```

;     REGISTRATION NUMBER: 28,678
;     REFERENCE/DOCKET NUMBER: 40028-A-PCT-US
;     TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (212) 278-0400
;     TELEFAX: (212) 391-0526
;     TELEX:
;     INFORMATION FOR SEQ ID NO: 20:
;     SEQUENCE CHARACTERISTICS:
;     LENGTH: 1170 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-313-288B-20

```

```

Query Match          91.7%; Score 55; DB 1; Length 1170;
Best Local Similarity 91.7%; Pred. No. 0.013;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 FQGVLLQVRFVF 12
        ||||| |||||
Db      208 FQGVLLQVRFVF 219

```

# RESULT 2

```

US-09-134-000C-3511
; Sequence 3511, Application US/09134000C
; Patent No. 6617156
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 032796-032
; CURRENT APPLICATION NUMBER: US/09/134,000C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/055,778
; PRIOR FILING DATE: 1997-08-15
; NUMBER OF SEQ ID NOS: 6812
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3511
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (29)..(70)
; OTHER INFORMATION: Amino acids 29, 30, 59 & 70 are Xaa wherein Xaa = any
amino acid.
US-09-134-000C-3511

```

```

Query Match          65.0%; Score 39; DB 4; Length 169;
Best Local Similarity 77.8%; Pred. No. 2.5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 FQGVLLQVR 9
        |||||:|:|
Db      76 FQGVLLRQIR 84

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:39:01 ; Search time 33.8326 Seconds  
 (without alignments)  
 93.153 Million cell updates/sec

Title: US-10-030-735-20  
 Perfect score: 60  
 Sequence: 1 FQGVLQQVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%				
No.	Score	Match Length	DB	ID		Description



1	55	91.7	240	12	US-10-419-462-40	Sequence 40, Appl
2	55	91.7	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	55	91.7	831	12	US-09-939-853A-97	Sequence 97, Appl
4	55	91.7	831	12	US-09-939-853A-98	Sequence 98, Appl
5	55	91.7	1152	9	US-09-919-603-1	Sequence 1, Appli
6	55	91.7	1170	12	US-10-211-462-38	Sequence 38, Appl
7	55	91.7	1170	12	US-10-231-956A-482	Sequence 482, App
8	55	91.7	1170	12	US-10-419-462-38	Sequence 38, Appl
9	55	91.7	1170	14	US-10-020-141-12	Sequence 12, Appl
10	55	91.7	1170	14	US-10-017-721-2	Sequence 2, Appli
11	55	91.7	1170	14	US-10-021-660-114	Sequence 114, App
12	55	91.7	1170	14	US-10-008-093-2	Sequence 2, Appli
13	55	91.7	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	39	65.0	15	15	US-10-285-394-153	Sequence 153, App
15	37	61.7	16	9	US-09-822-682-6	Sequence 6, Appli
16	37	61.7	151	9	US-09-925-299-1251	Sequence 1251, Ap
17	37	61.7	151	10	US-09-925-299-1251	Sequence 1251, Ap
18	37	61.7	1168	9	US-09-919-603-2	Sequence 2, Appli
19	37	61.7	1172	9	US-09-919-770-4	Sequence 4, Appli
20	37	61.7	1172	9	US-09-822-682-2	Sequence 2, Appli
21	37	61.7	1172	10	US-09-866-050A-658	Sequence 658, App
22	37	61.7	1172	14	US-10-060-036-171	Sequence 171, App
23	37	61.7	1172	14	US-10-020-141-14	Sequence 14, Appl
24	37	61.7	1172	14	US-10-017-724-2	Sequence 2, Appli
25	37	61.7	1172	14	US-10-301-822-203	Sequence 203, App
26	37	61.7	1172	14	US-10-008-093-4	Sequence 4, Appli
27	37	61.7	1172	14	US-10-007-781-2	Sequence 2, Appli
28	37	61.7	1172	15	US-10-372-683-16	Sequence 16, Appl
29	36	60.0	359	15	US-10-369-493-10609	Sequence 10609, A
30	36	60.0	459	12	US-10-282-122A-67080	Sequence 67080, A
31	35	58.3	75	12	US-10-424-599-271195	Sequence 271195,
32	35	58.3	83	12	US-10-424-599-150515	Sequence 150515,
33	35	58.3	1360	15	US-10-334-143-17	Sequence 17, Appl
34	34	56.7	106	12	US-10-424-599-242530	Sequence 242530,
35	34	56.7	139	15	US-10-108-260A-3323	Sequence 3323, Ap
36	34	56.7	363	15	US-10-080-334-84	Sequence 84, Appl
37	34	56.7	459	15	US-10-369-493-18328	Sequence 18328, A
38	34	56.7	490	15	US-10-080-334-250	Sequence 250, App
39	34	56.7	490	15	US-10-108-260A-4335	Sequence 4335, Ap
40	34	56.7	539	15	US-10-422-536-16	Sequence 16, Appl
41	34	56.7	565	15	US-10-369-493-6976	Sequence 6976, Ap
42	34	56.7	950	15	US-10-259-194A-58	Sequence 58, Appl
43	34	56.7	1247	9	US-09-815-242-4886	Sequence 4886, Ap
44	34	56.7	1271	9	US-09-815-242-10717	Sequence 10717, A
45	34	56.7	1271	12	US-10-282-122A-42498	Sequence 42498, A
46	34	56.7	1300	10	US-09-896-923-3	Sequence 3, Appli
47	34	56.7	1300	15	US-10-223-650-3	Sequence 3, Appli
48	34	56.7	1508	15	US-10-369-493-4104	Sequence 4104, Ap
49	34	56.7	1751	15	US-10-435-766-103	Sequence 103, App
50	33	55.0	51	12	US-10-424-599-198354	Sequence 198354,
51	33	55.0	69	12	US-10-424-599-206194	Sequence 206194,
52	33	55.0	125	12	US-10-424-599-259520	Sequence 259520,
53	33	55.0	200	12	US-10-282-122A-72332	Sequence 72332, A
54	33	55.0	276	12	US-10-335-977-5848	Sequence 5848, Ap
55	33	55.0	333	14	US-10-106-698-5027	Sequence 5027, Ap
56	33	55.0	336	9	US-09-925-300-1359	Sequence 1359, Ap

57	33	55.0	386	9	US-09-881-752A-74	Sequence 74, Appl
58	33	55.0	386	10	US-09-988-067B-32	Sequence 32, Appl
59	33	55.0	386	12	US-10-335-977-5849	Sequence 5849, Ap
60	33	55.0	394	12	US-10-335-977-5850	Sequence 5850, Ap
61	33	55.0	453	12	US-10-425-114-54475	Sequence 54475, A
62	33	55.0	453	15	US-10-094-749-1719	Sequence 1719, Ap
63	33	55.0	516	12	US-10-282-122A-77301	Sequence 77301, A
64	33	55.0	599	12	US-10-282-122A-48966	Sequence 48966, A
65	33	55.0	609	15	US-10-108-260A-3971	Sequence 3971, Ap
66	33	55.0	633	12	US-10-425-114-54254	Sequence 54254, A
67	33	55.0	639	12	US-10-425-114-54257	Sequence 54257, A
68	33	55.0	748	12	US-10-425-114-54253	Sequence 54253, A
69	33	55.0	815	15	US-10-394-322A-23	Sequence 23, Appl
70	33	55.0	861	12	US-10-425-114-54256	Sequence 54256, A
71	33	55.0	952	15	US-10-369-493-13682	Sequence 13682, A
72	33	55.0	981	14	US-10-128-714-8364	Sequence 8364, Ap
73	33	55.0	1000	14	US-10-128-714-3364	Sequence 3364, Ap
74	33	55.0	1224	12	US-10-282-122A-70506	Sequence 70506, A
75	33	55.0	1240	12	US-10-282-122A-57585	Sequence 57585, A
76	33	55.0	1835	15	US-10-369-493-1489	Sequence 1489, Ap
77	32.5	54.2	475	15	US-10-369-493-13733	Sequence 13733, A
78	32	53.3	42	12	US-10-424-599-146769	Sequence 146769,
79	32	53.3	81	14	US-10-029-386-29347	Sequence 29347, A
80	32	53.3	89	12	US-10-424-599-153242	Sequence 153242,
81	32	53.3	101	12	US-10-424-599-161904	Sequence 161904,
82	32	53.3	186	12	US-10-282-122A-48526	Sequence 48526, A
83	32	53.3	190	14	US-10-029-386-32211	Sequence 32211, A
84	32	53.3	214	9	US-09-747-155-383	Sequence 383, App
85	32	53.3	215	14	US-10-029-386-33750	Sequence 33750, A
86	32	53.3	226	9	US-09-925-300-1217	Sequence 1217, Ap
87	32	53.3	226	14	US-10-359-385-3	Sequence 3, Appli
88	32	53.3	241	12	US-10-282-122A-55154	Sequence 55154, A
89	32	53.3	280	14	US-10-156-761-12005	Sequence 12005, A
90	32	53.3	312	9	US-09-886-055-511	Sequence 511, App
91	32	53.3	312	10	US-09-804-291-511	Sequence 511, App
92	32	53.3	312	12	US-10-433-581-7	Sequence 7, Appli
93	32	53.3	312	15	US-10-085-198-198	Sequence 198, App
94	32	53.3	319	15	US-10-085-198-204	Sequence 204, App
95	32	53.3	329	12	US-10-424-599-177504	Sequence 177504,
96	32	53.3	359	14	US-10-017-161-678	Sequence 678, App
97	32	53.3	359	15	US-10-292-798-594	Sequence 594, App
98	32	53.3	368	10	US-09-819-607-5	Sequence 5, Appli
99	32	53.3	368	15	US-10-074-978A-156	Sequence 156, App
100	32	53.3	388	12	US-10-425-114-67593	Sequence 67593, A
101	32	53.3	390	14	US-10-420-845-6	Sequence 6, Appli
102	32	53.3	413	15	US-10-369-493-16902	Sequence 16902, A
103	32	53.3	493	12	US-10-425-114-63801	Sequence 63801, A
104	32	53.3	510	14	US-10-311-879-22	Sequence 22, Appl
105	32	53.3	535	12	US-10-282-122A-57701	Sequence 57701, A
106	32	53.3	557	9	US-09-815-242-13521	Sequence 13521, A
107	32	53.3	559	14	US-10-284-499-47	Sequence 47, Appl
108	32	53.3	601	9	US-09-815-242-12358	Sequence 12358, A
109	32	53.3	602	9	US-09-815-242-10802	Sequence 10802, A
110	32	53.3	604	9	US-09-815-242-5260	Sequence 5260, Ap
111	32	53.3	611	12	US-10-282-122A-56907	Sequence 56907, A
112	32	53.3	612	12	US-10-282-122A-60475	Sequence 60475, A
113	32	53.3	613	12	US-10-282-122A-74655	Sequence 74655, A

114	32	53.3	614	12	US-10-282-122A-71407	Sequence 71407, A
115	32	53.3	614	12	US-10-282-122A-72037	Sequence 72037, A
116	32	53.3	615	12	US-10-282-122A-44478	Sequence 44478, A
117	32	53.3	615	12	US-10-282-122A-70698	Sequence 70698, A
118	32	53.3	620	12	US-10-282-122A-73864	Sequence 73864, A
119	32	53.3	634	12	US-10-282-122A-46137	Sequence 46137, A
120	32	53.3	651	14	US-10-420-845-5	Sequence 5, Appli
121	32	53.3	763	15	US-10-108-260A-3277	Sequence 3277, Ap
122	32	53.3	805	12	US-10-112-944-346	Sequence 346, App
123	32	53.3	807	15	US-10-369-493-3780	Sequence 3780, Ap
124	32	53.3	1070	14	US-10-420-845-22	Sequence 22, Appl
125	32	53.3	1207	9	US-09-064-199-16	Sequence 16, Appl
126	32	53.3	1213	9	US-09-815-242-5358	Sequence 5358, Ap
127	32	53.3	1217	9	US-09-815-242-12523	Sequence 12523, A
128	32	53.3	1217	12	US-10-282-122A-43796	Sequence 43796, A
129	32	53.3	1228	12	US-10-282-122A-71769	Sequence 71769, A
130	32	53.3	1317	9	US-09-815-242-5118	Sequence 5118, Ap
131	32	53.3	1317	12	US-10-282-122A-43495	Sequence 43495, A
132	32	53.3	1504	9	US-09-932-145-7	Sequence 7, Appli
133	32	53.3	1504	14	US-10-284-499-1	Sequence 1, Appli
134	32	53.3	1504	15	US-10-395-812-7	Sequence 7, Appli
135	32	53.3	3500	14	US-10-153-219-2	Sequence 2, Appli
136	32	53.3	3537	14	US-10-153-219-15	Sequence 15, Appl
137	31	51.7	50	9	US-09-971-980-64	Sequence 64, Appl
138	31	51.7	50	12	US-10-424-599-187405	Sequence 187405,
139	31	51.7	50	12	US-10-424-599-193492	Sequence 193492,
140	31	51.7	53	11	US-09-864-408A-6976	Sequence 6976, Ap
141	31	51.7	59	12	US-10-424-599-177312	Sequence 177312,
142	31	51.7	64	9	US-09-971-980-62	Sequence 62, Appl
143	31	51.7	75	12	US-10-424-599-165575	Sequence 165575,
144	31	51.7	81	12	US-10-424-599-209278	Sequence 209278,
145	31	51.7	82	12	US-10-424-599-243633	Sequence 243633,
146	31	51.7	108	12	US-10-282-122A-67175	Sequence 67175, A
147	31	51.7	113	14	US-10-055-882-4	Sequence 4, Appli
148	31	51.7	121	12	US-10-425-114-67826	Sequence 67826, A
149	31	51.7	130	9	US-09-949-192-41	Sequence 41, Appl
150	31	51.7	140	12	US-10-335-977-6317	Sequence 6317, Ap

#### ALIGNMENTS

##### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 91.7%; Score 55; DB 12; Length 240;  
Best Local Similarity 91.7%; Pred. No. 0.023;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLQQVRFVF 12  
||| ||| ||| |||  
Db 190 FQGVLLQNVRVFVF 201

#### RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 91.7%; Score 55; DB 9; Length 466;  
Best Local Similarity 91.7%; Pred. No. 0.046;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLQQVRFVF 12  
||| ||| ||| |||  
Db 261 FQGVLLQNVRVFVF 272

#### RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding Same

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:31:01 ; Search time 6.66079 Seconds  
 (without alignments)  
 35.719 Million cell updates/sec

Title: US-10-030-735-20  
 Perfect score: 60  
 Sequence: 1 FQGVLLQQVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 167301 seqs, 19826498 residues

Total number of hits satisfying chosen parameters: 167301

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Pending\_Patents\_AA\_New:\*  
 1: /cgn2\_6/ptodata/2/paa/PCT\_NEW\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
 7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	55	91.7	1170	7	US-60-539-605-28	Sequence 28, Appl
2	36	60.0	517	6	US-10-761-781-1	Sequence 1, Appli
3	33	55.0	386	6	US-10-786-892-74	Sequence 74, Appl
4	33	55.0	533	1	PCT-US04-05092-41	Sequence 41, Appl
5	33	55.0	1242	6	US-10-417-884A-5241	Sequence 5241, Ap
6	32	53.3	92	1	PCT-US04-07096-561	Sequence 561, App
7	32	53.3	190	6	US-10-767-701-55109	Sequence 55109, A
8	32	53.3	282	6	US-10-767-701-38944	Sequence 38944, A
9	32	53.3	507	1	PCT-US04-07096-566	Sequence 566, App
10	32	53.3	624	6	US-10-417-884A-5745	Sequence 5745, Ap

11	32	53.3	3500	1	PCT-US03-37356-2	Sequence 2, Appli
12	32	53.3	3537	1	PCT-US03-37356-15	Sequence 15, Appl
13	31	51.7	102	1	PCT-US03-32645-30	Sequence 30, Appl
14	31	51.7	108	6	US-10-767-701-56318	Sequence 56318, A
15	31	51.7	113	6	US-10-487-078-17	Sequence 17, Appl
16	31	51.7	198	6	US-10-767-701-56675	Sequence 56675, A
17	31	51.7	208	1	PCT-US04-04499-104	Sequence 104, App
18	31	51.7	218	1	PCT-US04-04499-103	Sequence 103, App
19	31	51.7	220	6	US-10-786-892-48	Sequence 48, Appl
20	31	51.7	328	6	US-10-767-701-38418	Sequence 38418, A
21	31	51.7	329	1	PCT-US04-04499-35	Sequence 35, Appl
22	31	51.7	329	1	PCT-US04-04499-53	Sequence 53, Appl
23	31	51.7	329	1	PCT-US04-04499-178	Sequence 178, App
24	31	51.7	329	1	PCT-US04-04499-250	Sequence 250, App
25	31	51.7	370	6	US-10-784-480-210	Sequence 210, App
26	31	51.7	445	6	US-10-417-884A-5506	Sequence 5506, Ap
27	31	51.7	578	7	US-60-548-091-513	Sequence 513, App
28	31	51.7	810	1	PCT-US04-03808-2	Sequence 2, Appli
29	31	51.7	810	1	PCT-US04-03808-4	Sequence 4, Appli
30	30	50.0	144	6	US-10-767-701-48948	Sequence 48948, A
31	30	50.0	148	6	US-10-100-683-7434	Sequence 7434, Ap
32	30	50.0	149	6	US-10-767-701-57905	Sequence 57905, A
33	30	50.0	200	6	US-10-767-701-37410	Sequence 37410, A
34	30	50.0	203	1	PCT-US04-07716-2	Sequence 2, Appli
35	30	50.0	209	1	PCT-US04-07716-5	Sequence 5, Appli
36	30	50.0	216	6	US-10-781-014-420	Sequence 420, App
37	30	50.0	226	6	US-10-784-480-362	Sequence 362, App
38	30	50.0	240	1	PCT-US04-07716-23	Sequence 23, Appl
39	30	50.0	246	6	US-10-013-910A-338	Sequence 338, App
40	30	50.0	278	6	US-10-490-318-18	Sequence 18, Appl
41	30	50.0	309	6	US-10-100-683-9072	Sequence 9072, Ap
42	30	50.0	362	6	US-10-100-683-7603	Sequence 7603, Ap
43	30	50.0	365	6	US-10-775-965-8	Sequence 8, Appli
44	30	50.0	365	6	US-10-668-035-27	Sequence 27, Appl
45	30	50.0	428	6	US-10-796-907-28	Sequence 28, Appl
46	30	50.0	428	6	US-10-796-907-44	Sequence 44, Appl
47	30	50.0	487	6	US-10-781-014-418	Sequence 418, App
48	30	50.0	548	6	US-10-767-701-44071	Sequence 44071, A
49	29	48.3	14	6	US-10-111-983-4131	Sequence 4131, Ap
50	29	48.3	14	6	US-10-111-983-18000	Sequence 18000, A
51	29	48.3	14	6	US-10-111-983-30046	Sequence 30046, A
52	29	48.3	15	6	US-10-111-983-4122	Sequence 4122, Ap
53	29	48.3	15	6	US-10-111-983-17991	Sequence 17991, A
54	29	48.3	15	6	US-10-111-983-30040	Sequence 30040, A
55	29	48.3	40	6	US-10-100-683-8020	Sequence 8020, Ap
56	29	48.3	69	6	US-10-793-479-7425	Sequence 7425, Ap
57	29	48.3	98	5	US-09-796-692B-2345	Sequence 2345, Ap
58	29	48.3	103	6	US-10-767-701-57253	Sequence 57253, A
59	29	48.3	108	6	US-10-767-701-38744	Sequence 38744, A
60	29	48.3	116	6	US-10-767-701-38126	Sequence 38126, A
61	29	48.3	134	6	US-10-767-701-63040	Sequence 63040, A
62	29	48.3	192	6	US-10-767-701-40959	Sequence 40959, A
63	29	48.3	215	6	US-10-767-701-41591	Sequence 41591, A
64	29	48.3	270	6	US-10-767-701-39482	Sequence 39482, A
65	29	48.3	278	6	US-10-767-701-44161	Sequence 44161, A
66	29	48.3	282	6	US-10-767-701-38931	Sequence 38931, A
67	29	48.3	306	1	PCT-US04-04499-266	Sequence 266, App

68	29	48.3	315	6	US-10-767-701-41601	Sequence 41601, A
69	29	48.3	331	6	US-10-664-421-91	Sequence 91, Appl
70	29	48.3	338	6	US-10-767-701-41704	Sequence 41704, A
71	29	48.3	345	6	US-10-796-907-60	Sequence 60, Appl
72	29	48.3	345	6	US-10-796-907-70	Sequence 70, Appl
73	29	48.3	345	6	US-10-796-907-86	Sequence 86, Appl
74	29	48.3	345	6	US-10-796-907-138	Sequence 138, App
75	29	48.3	366	6	US-10-767-701-40407	Sequence 40407, A
76	29	48.3	570	7	US-60-551-161-68	Sequence 68, Appl
77	29	48.3	570	7	US-60-551-121-108	Sequence 108, App
78	29	48.3	586	1	PCT-US04-07412-796	Sequence 796, App
79	29	48.3	598	6	US-10-761-033-12	Sequence 12, Appl
80	29	48.3	605	6	US-10-761-033-8	Sequence 8, Appli
81	29	48.3	621	6	US-10-761-033-4	Sequence 4, Appli
82	29	48.3	800	6	US-10-767-701-45770	Sequence 45770, A
83	29	48.3	816	1	PCT-US04-03808-8	Sequence 8, Appli
84	29	48.3	816	1	PCT-US04-03808-10	Sequence 10, Appl
85	29	48.3	816	1	PCT-US04-03808-14	Sequence 14, Appl
86	28	46.7	51	6	US-10-767-701-62876	Sequence 62876, A
87	28	46.7	51	6	US-10-100-683-6285	Sequence 6285, Ap
88	28	46.7	51	6	US-10-800-834-184	Sequence 184, App
89	28	46.7	60	6	US-10-767-701-55517	Sequence 55517, A
90	28	46.7	68	6	US-10-807-837-48	Sequence 48, Appl
91	28	46.7	68	6	US-10-807-997-48	Sequence 48, Appl
92	28	46.7	71	6	US-10-767-701-47117	Sequence 47117, A
93	28	46.7	72	6	US-10-800-834-407	Sequence 407, App
94	28	46.7	107	6	US-10-767-701-43690	Sequence 43690, A
95	28	46.7	108	6	US-10-767-701-53440	Sequence 53440, A
96	28	46.7	119	6	US-10-100-683-10125	Sequence 10125, A
97	28	46.7	138	6	US-10-767-701-43423	Sequence 43423, A
98	28	46.7	148	6	US-10-767-701-53447	Sequence 53447, A
99	28	46.7	167	6	US-10-767-701-38208	Sequence 38208, A
100	28	46.7	169	6	US-10-487-578-8	Sequence 8, Appli
101	28	46.7	175	6	US-10-767-701-58182	Sequence 58182, A
102	28	46.7	187	6	US-10-767-701-35834	Sequence 35834, A
103	28	46.7	197	6	US-10-767-701-61444	Sequence 61444, A
104	28	46.7	200	6	US-10-767-701-57238	Sequence 57238, A
105	28	46.7	212	6	US-10-807-837-62	Sequence 62, Appl
106	28	46.7	212	6	US-10-807-997-62	Sequence 62, Appl
107	28	46.7	226	6	US-10-784-480-64	Sequence 64, Appl
108	28	46.7	262	6	US-10-767-701-47160	Sequence 47160, A
109	28	46.7	271	1	PCT-US04-07467-653	Sequence 653, App
110	28	46.7	286	6	US-10-767-701-45014	Sequence 45014, A
111	28	46.7	299	6	US-10-417-884A-5981	Sequence 5981, Ap
112	28	46.7	309	6	US-10-417-884A-6518	Sequence 6518, Ap
113	28	46.7	312	6	US-10-767-701-41715	Sequence 41715, A
114	28	46.7	328	6	US-10-417-884A-6973	Sequence 6973, Ap
115	28	46.7	334	6	US-10-767-701-43860	Sequence 43860, A
116	28	46.7	360	6	US-10-417-884A-6666	Sequence 6666, Ap
117	28	46.7	365	6	US-10-784-480-264	Sequence 264, App
118	28	46.7	392	6	US-10-767-701-41232	Sequence 41232, A
119	28	46.7	419	6	US-10-796-307-663	Sequence 663, App
120	28	46.7	419	7	US-60-550-051-293	Sequence 293, App
121	28	46.7	436	5	US-09-979-167-153	Sequence 153, App
122	28	46.7	445	1	PCT-US04-05654-2896	Sequence 2896, Ap
123	28	46.7	461	6	US-10-490-594-10	Sequence 10, Appl
124	28	46.7	466	6	US-10-775-204-2219	Sequence 2219, Ap

125	28	46.7	466	6	US-10-796-307-662	Sequence 662, App
126	28	46.7	466	7	US-60-550-051-292	Sequence 292, App
127	28	46.7	467	6	US-10-486-805-25	Sequence 25, Appl
128	28	46.7	483	6	US-10-781-014-762	Sequence 762, App
129	28	46.7	490	6	US-10-486-805-26	Sequence 26, Appl
130	28	46.7	492	6	US-10-100-683-10905	Sequence 10905, A
131	28	46.7	580	6	US-10-796-280-808	Sequence 808, App
132	28	46.7	581	6	US-10-807-837-42	Sequence 42, Appl
133	28	46.7	581	6	US-10-807-997-42	Sequence 42, Appl
134	28	46.7	587	6	US-10-417-884A-6405	Sequence 6405, Ap
135	28	46.7	659	5	US-09-980-437A-28	Sequence 28, Appl
136	28	46.7	724	6	US-10-416-898-18	Sequence 18, Appl
137	28	46.7	903	6	US-10-489-740-145	Sequence 145, App
138	28	46.7	1098	1	PCT-US04-07268-221	Sequence 221, App
139	28	46.7	1098	6	US-10-788-792-221	Sequence 221, App
140	28	46.7	1112	6	US-10-398-186-22	Sequence 22, Appl
141	28	46.7	1304	6	US-10-489-740-157	Sequence 157, App
142	28	46.7	1322	6	US-10-796-280-1082	Sequence 1082, Ap
143	28	46.7	1333	6	US-10-796-280-1076	Sequence 1076, Ap
144	28	46.7	1437	7	US-60-551-161-96	Sequence 96, Appl
145	28	46.7	1868	6	US-10-796-280-1078	Sequence 1078, Ap
146	28	46.7	1922	6	US-10-796-280-1080	Sequence 1080, Ap
147	28	46.7	1954	6	US-10-796-280-1081	Sequence 1081, Ap
148	28	46.7	1979	6	US-10-796-280-1079	Sequence 1079, Ap
149	28	46.7	6306	1	PCT-US04-09049-5	Sequence 5, Appli
150	28	46.7	6893	5	US-09-976-059A-14	Sequence 14, Appl

#### ALIGNMENTS

##### RESULT 1

US-60-539-605-28

; Sequence 28, Application US/60539605

; GENERAL INFORMATION:

; APPLICANT: Boyle, Bryan

; APPLICANT: Funk, Walter

; APPLICANT: Kakitani, Makoto

; APPLICANT: Oshima, Takeshi

; APPLICANT: Park, Eun Ju

; APPLICANT: Tang, Y. Tom

; APPLICANT: Tomizuka, Kazuma

; APPLICANT: Yagi, Mikio

; TITLE OF INVENTION: Gastrointestinal Proliferative Factor and uses thereof

; FILE REFERENCE: NUVO-12

; CURRENT APPLICATION NUMBER: US/60/539,605

; CURRENT FILING DATE: 2004-01-27

; NUMBER OF SEQ ID NOS: 67

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 28

; LENGTH: 1170

; TYPE: PRT

; ORGANISM: Homo sapiens

US-60-539-605-28

Query Match	91.7%;	Score 55;	DB 7;	Length 1170;
Best Local Similarity	91.7%;	Pred. No. 0.0063;		



```

Matches    11;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 FQGVLLQQVRFVF 12
           ||||| |||||
Db          208 FQGVLLQNVRVFVF 219

```

RESULT 2

```

US-10-761-781-1
; Sequence 1, Application US/10761781
; GENERAL INFORMATION:
; APPLICANT: Vale, Ronald
; APPLICANT: Hartman, James
; TITLE OF INVENTION: Assays for the Detection of Microtubule Depolymerization
Inhibitors
; FILE REFERENCE: UCSD-04765
; CURRENT APPLICATION NUMBER: US/10/761,781
; CURRENT FILING DATE: 2004-01-20
; PRIOR APPLICATION NUMBER: US/09/673,222A
; PRIOR FILING DATE: 2000-12-04
; PRIOR APPLICATION NUMBER: PCT/US99/08086
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/081,734
; PRIOR FILING DATE: 1998-04-14
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Strongylocentrotus purpuratus
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: katanin p60 subunit
US-10-761-781-1

```

```

Query Match          60.0%;  Score 36;  DB 6;  Length 517;
Best Local Similarity 66.7%;  Pred. No. 16;
Matches    6;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

```

```

Qy          1 FQGVLLQQVR 9
           :|||||::
Db          29 YQGVLLQQIQ 37

```

RESULT 3

```

US-10-786-892-74
; Sequence 74, Application US/10786892
; GENERAL INFORMATION:
; APPLICANT: Kleanthous, Harold
; APPLICANT: Al-Garawi, Amal
; APPLICANT: Miller, Charles
; APPLICANT: Tomb, Jean-Francois
; APPLICANT: Oomen, Raymond P.
; TITLE OF INVENTION: Identification of Polynucleotides
; TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the
Helicobacter
; TITLE OF INVENTION: Genome

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:03:50 ; Search time 9.67401 Seconds  
 (without alignments)  
 119.320 Million cell updates/sec

Title: US-10-030-735-20  
 Perfect score: 60  
 Sequence: 1 FQGVLLQQVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	55	91.7	229	2	S57957	thrombospondin 1 -	
2	55	91.7	1170	1	TSHUP1	thrombospondin 1 p	
3	55	91.7	1170	2	A40558	thrombospondin 1 p	
4	38	63.3	212	2	D81929	probable imidazole	
5	38	63.3	469	2	T33595	hypothetical prote	
6	37	61.7	151	2	C57253	tRNA-pseudouridine	
7	37	61.7	1172	1	TSHUP2	thrombospondin 2 p	
8	37	61.7	1172	2	A42587	thrombospondin 2 p	
9	36	60.0	102	2	D81910	hypothetical prote	
10	36	60.0	145	2	T20985	hypothetical prote	
11	36	60.0	162	2	B88349	protein F15D4.3 [i	
12	36	60.0	1299	2	AH2090	two-component hybr	
13	35	58.3	102	2	G81105	hypothetical prote	

14	35	58.3	313	2	S51728	NSP3 protein - hum
15	35	58.3	471	2	S45068	53K glycoprotein -
16	35	58.3	516	2	E84019	methylmalonyl-CoA
17	35	58.3	527	2	T49212	hypothetical prote
18	35	58.3	893	2	T36795	probable penicilli
19	35	58.3	1142	1	GNVUPH	M polyprotein prec
20	35	58.3	1148	1	JQ1604	M polyprotein prec
21	35	58.3	1178	1	A39804	thrombospondin pre
22	35	58.3	1333	1	XOHUDH	xanthine dehydroge
23	34	56.7	165	2	C83862	hypothetical prote
24	34	56.7	175	2	S75258	hypothetical prote
25	34	56.7	295	2	B54575	35K GTP-binding pr
26	34	56.7	311	2	F70184	ribose/galactose A
27	34	56.7	400	2	T24890	hypothetical prote
28	34	56.7	402	2	B82418	hypothetical prote
29	34	56.7	459	2	D86669	amino acid permeas
30	34	56.7	477	1	P2WLEP	L2 protein - Europ
31	34	56.7	505	2	E87021	probable integral-
32	34	56.7	552	2	A70709	probable ptrBa pro
33	34	56.7	565	2	T23843	hypothetical prote
34	34	56.7	683	2	AC0741	oligopeptidase B (
35	34	56.7	686	1	E64946	oligopeptidase B (
36	34	56.7	686	2	C90948	proteinase II [imp
37	34	56.7	686	2	G85796	proteinase II [imp
38	34	56.7	724	2	F87187	proteinase II [imp
39	34	56.7	747	2	AE2929	two component resp
40	34	56.7	783	2	A98353	probable transcrip
41	34	56.7	913	2	T52485	neurofilament prot
42	34	56.7	1142	2	T39103	probable negative
43	34	56.7	1312	2	S68593	DNA-directed DNA p
44	34	56.7	1670	2	S71551	DNA-directed DNA p
45	33	55.0	82	2	F96625	hypothetical prote
46	33	55.0	106	2	C90261	hypothetical prote
47	33	55.0	190	2	A37840	grpE protein homol
48	33	55.0	190	2	G81676	grpE protein TC067
49	33	55.0	194	2	T04180	ribosomal protein
50	33	55.0	269	2	A49334	Ras homolog Rad -
51	33	55.0	275	1	WMBEH4	UL34 protein - hum
52	33	55.0	295	2	E72462	probable lipoic ac
53	33	55.0	295	2	I49117	Ras-like protein K
54	33	55.0	296	2	A54575	35K GTP-binding pr
55	33	55.0	296	2	I38745	kinase-inducible r
56	33	55.0	313	2	S51769	nonstructural prot
57	33	55.0	323	2	AH0297	periplasmic-bindin
58	33	55.0	371	2	B69451	conserved hypothet
59	33	55.0	386	2	D71851	tetracycline resis
60	33	55.0	386	2	E64665	tetracycline resis
61	33	55.0	401	2	E84757	hypothetical prote
62	33	55.0	439	2	E36840	C14L protein - var
63	33	55.0	439	2	H72154	E10L protein - var
64	33	55.0	439	2	T30787	hypothetical prote
65	33	55.0	439	2	F42507	F10L protein - vac
66	33	55.0	439	2	T28472	hypothetical prote
67	33	55.0	455	2	D83264	hypothetical prote
68	33	55.0	456	2	H97735	hypothetical prote
69	33	55.0	465	2	AC0347	probable membrane
70	33	55.0	516	2	G82182	probable NADH dehy

71	33	55.0	582	2	T07953	lectin-like protei
72	33	55.0	582	2	T07952	lectin-like protei
73	33	55.0	633	2	H95932	probable adenylyl-
74	33	55.0	641	1	ZZZRQ	adenylyl-sulfate k
75	33	55.0	641	2	E95320	adenylyl-sulfate k
76	33	55.0	715	2	S70397	zona pellucida gly
77	33	55.0	815	2	B56708	extracellular sign
78	33	55.0	1011	2	C84524	probable disease r
79	33	55.0	1148	1	GNVUNE	M polyprotein prec
80	33	55.0	1331	1	XORTDH	xanthine dehydroge
81	33	55.0	1335	1	XOMSDH	xanthine dehydroge
82	33	55.0	1493	2	A44224	DNA repair helicase
83	33	55.0	1835	2	S46082	urea carboxylase (
84	33	55.0	1995	2	T08166	probable membrane
85	32	53.3	111	2	S52596	probable membrane
86	32	53.3	182	2	A75370	hypothetical prote
87	32	53.3	190	2	A71521	probable hsp-70 co
88	32	53.3	241	2	F71558	probable tRNA pseu
89	32	53.3	241	2	F81709	tRNA pseudouridine
90	32	53.3	260	2	G82166	smtA protein VC171
91	32	53.3	274	2	AC3066	conserved hypothet
92	32	53.3	277	2	JQ1288	regulatory protein
93	32	53.3	278	2	AF0944	L-rhamnose operon
94	32	53.3	287	2	F98220	hypothetical prote
95	32	53.3	298	2	C87403	FdhD protein [impo
96	32	53.3	313	2	S51729	NSP3 protein - hum
97	32	53.3	326	2	H83832	hypothetical prote
98	32	53.3	342	2	S76680	hypothetical prote
99	32	53.3	380	2	T26165	hypothetical prote
100	32	53.3	394	2	AH0362	nucleoside permeas
101	32	53.3	413	2	B87451	membrane protein,
102	32	53.3	417	2	H83708	hypothetical prote
103	32	53.3	467	2	D84938	H+-transporting tw
104	32	53.3	487	1	H81109	IMP dehydrogenase
105	32	53.3	487	2	F81906	IMP dehydrogenase
106	32	53.3	511	2	S77350	hypothetical prote
107	32	53.3	612	2	AC1208	GTP-binding elonga
108	32	53.3	612	2	AF1564	GTP-binding elonga
109	32	53.3	613	2	G86876	GTP-binding protei
110	32	53.3	613	2	A95079	elongation factor
111	32	53.3	615	2	B89881	hypothetical prote
112	32	53.3	620	2	F97946	hypothetical prote
113	32	53.3	661	2	T32518	hypothetical prote
114	32	53.3	682	2	AG3351	transporter BMEI07
115	32	53.3	896	2	A41273	glutamate receptor
116	32	53.3	906	2	A40222	glutamate receptor
117	32	53.3	907	1	ACRTK1	glutamate receptor
118	32	53.3	907	2	A40170	glutamate receptor
119	32	53.3	907	2	S12874	glutamate receptor
120	32	53.3	915	2	T21147	hypothetical prote
121	32	53.3	993	2	T17230	hypothetical prote
122	32	53.3	1177	2	I64233	hypothetical prote
123	32	53.3	1217	2	H89863	hypothetical prote
124	32	53.3	1317	2	F83310	conserved hypothet
125	32	53.3	1702	2	S42459	DNA-directed DNA p
126	32	53.3	1838	1	CGHU1V	collagen alpha 1(V
127	32	53.3	1843	2	S18803	collagen alpha 1(V

128	32	53.3	15281	2	S41309	cyclosporin synthe
129	31	51.7	62	2	D69894	hypothetical prote
130	31	51.7	157	2	T46440	hypothetical prote
131	31	51.7	211	2	C72509	hypothetical prote
132	31	51.7	220	2	H71824	hypothetical prote
133	31	51.7	220	2	B64692	competence lipopro
134	31	51.7	224	2	A65113	hypothetical 25.3
135	31	51.7	225	2	AB2374	ABC transporter AT
136	31	51.7	244	2	AC1765	B. subtilis TagA p
137	31	51.7	255	2	E75501	probable RNA methy
138	31	51.7	258	2	A72624	probable proteasom
139	31	51.7	277	2	B83856	4-hydroxybenzoate
140	31	51.7	277	2	F71335	probable A/G-speci
141	31	51.7	285	2	H95969	hypothetical prote
142	31	51.7	364	2	F81843	conserved hypothet
143	31	51.7	364	2	B81098	conserved hypothet
144	31	51.7	365	2	AF3217	DNA-damage-inducib
145	31	51.7	379	2	S09791	hypothetical prote
146	31	51.7	385	2	S56224	hypothetical prote
147	31	51.7	389	2	S33667	probable integrase
148	31	51.7	393	2	T35894	hypothetical prote
149	31	51.7	397	2	G64703	type IIS restricti
150	31	51.7	406	2	G71816	probable type II r

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 91.7%; Score 55; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0016;

Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	FQGV	LQQV	RFVF	12
Db	190	FQGV	LQNV	RFVF	201

##### RESULT 2

TSHUP1  
 thrombospondin 1 precursor - human  
 C;Species: Homo sapiens (man)  
 C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000  
 C;Accession: A26155; A34274; A30140; A25812; A05172; A42927  
 R;Lawler, J.; Hynes, R.O.  
 J. Cell Biol. 103, 1635-1648, 1986  
 A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.  
 A;Reference number: A26155; MUID:87057617; PMID:2430973  
 A;Accession: A26155  
 A;Molecule type: mRNA  
 A;Residues: 1-1170 <LAW>  
 A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138  
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing  
 R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.  
 J. Biol. Chem. 264, 11222-11227, 1989  
 A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.  
 A;Reference number: A34274; MUID:89291870; PMID:2544587  
 A;Accession: A34274  
 A;Molecule type: DNA  
 A;Residues: 1-166 <LAH>  
 A;Cross-references: GB:J04835  
 R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.  
 J. Cell Biol. 108, 729-736, 1989  
 A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.  
 A;Reference number: A30140; MUID:89139590; PMID:2918029  
 A;Accession: A30140  
 A;Molecule type: mRNA  
 A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>  
 A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465  
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing  
 R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.  
 Biochemistry 25, 8418-8425, 1986  
 A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.  
 A;Reference number: A25812; MUID:87157592; PMID:3030396  
 A;Accession: A25812  
 A;Molecule type: mRNA  
 A;Residues: 1-83,'A',85-397 <KOB>  
 A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354  
 R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.  
 Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986  
 A;Reference number: A05172; MUID:86287276; PMID:3461443  
 A;Accession: A05172  
 A;Molecule type: mRNA  
 A;Residues: 1-83,'A',85-374,'RC' <DIX>  
 A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801  
 A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing  
 R;Sun, X.; Skorstengaard, K.; Mosher, D.F.  
 J. Cell Biol. 118, 693-701, 1992

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:40:26 ; Search time 12.8831 Seconds  
(without alignments)  
48.087 Million cell updates/sec

Title: US-10-030-735-21  
Perfect score: 59  
Sequence: 1 FQGVLSVRFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result		%	Query				
No.	Score	Match	Length	DB	ID	Description	
1	56	94.9	1170	1	US-08-313-288B-20	Sequence 20, Appl	
2	38	64.4	175	4	US-09-252-991A-18826	Sequence 18826, A	
3	38	64.4	1172	1	US-08-313-288B-19	Sequence 19, Appl	
4	37	62.7	715	2	US-08-484-993B-10	Sequence 10, Appl	
5	37	62.7	715	2	US-08-484-158B-10	Sequence 10, Appl	
6	37	62.7	715	2	US-08-484-596A-10	Sequence 10, Appl	
7	37	62.7	715	2	US-08-480-150A-10	Sequence 10, Appl	
8	37	62.7	715	3	US-08-458-731-10	Sequence 10, Appl	
9	37	62.7	715	3	US-08-149-223A-10	Sequence 10, Appl	
10	36	61.0	390	3	US-08-650-766-7	Sequence 7, Appli	
11	36	61.0	390	3	US-08-922-635-6	Sequence 6, Appli	

12	36	61.0	390	4	US-09-389-487-7	Sequence 7, Appli
13	36	61.0	539	4	US-09-800-170-16	Sequence 16, Appl
14	36	61.0	559	4	US-09-364-206-47	Sequence 47, Appl
15	36	61.0	651	3	US-08-650-766-6	Sequence 6, Appli
16	36	61.0	651	3	US-08-922-635-5	Sequence 5, Appli
17	36	61.0	651	4	US-09-389-487-6	Sequence 6, Appli
18	36	61.0	1019	1	US-08-271-364A-7	Sequence 7, Appli
19	36	61.0	1019	2	US-08-222-715B-26	Sequence 26, Appl
20	36	61.0	1070	3	US-08-922-635-22	Sequence 22, Appl
21	36	61.0	1504	4	US-09-364-206-2	Sequence 2, Appli
22	35	59.3	440	1	US-08-307-499-15	Sequence 15, Appl
23	35	59.3	440	3	US-09-299-268-15	Sequence 15, Appl
24	35	59.3	944	4	US-09-134-000C-5578	Sequence 5578, Ap
25	34	57.6	169	4	US-09-134-000C-3511	Sequence 3511, Ap
26	34	57.6	1022	1	US-08-271-364A-8	Sequence 8, Appli
27	34	57.6	1022	2	US-08-222-715B-27	Sequence 27, Appl
28	34	57.6	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
29	34	57.6	1242	4	US-09-107-532A-5241	Sequence 5241, Ap
30	33	55.9	153	2	US-08-387-942C-51	Sequence 51, Appl
31	33	55.9	432	4	US-09-489-039A-7909	Sequence 7909, Ap
32	33	55.9	439	1	US-08-507-431-4	Sequence 4, Appli
33	33	55.9	439	2	US-08-700-546-2	Sequence 2, Appli
34	33	55.9	439	2	US-08-902-655A-4	Sequence 4, Appli
35	33	55.9	439	3	US-09-116-622-4	Sequence 4, Appli
36	33	55.9	439	3	US-09-219-277-4	Sequence 4, Appli
37	33	55.9	439	3	US-09-599-661-4	Sequence 4, Appli
38	33	55.9	872	2	US-08-387-942C-5	Sequence 5, Appli
39	32	54.2	153	2	US-08-387-942C-43	Sequence 43, Appl
40	32	54.2	159	4	US-09-149-476-499	Sequence 499, App
41	32	54.2	181	3	US-09-305-640-4	Sequence 4, Appli
42	32	54.2	207	4	US-08-811-519-24	Sequence 24, Appl
43	32	54.2	325	4	US-09-328-352-5100	Sequence 5100, Ap
44	32	54.2	365	2	US-08-724-974A-3	Sequence 3, Appli
45	32	54.2	365	4	US-09-170-496D-118	Sequence 118, App
46	32	54.2	365	4	US-09-170-496D-226	Sequence 226, App
47	32	54.2	365	4	US-09-364-425B-27	Sequence 27, Appl
48	32	54.2	445	4	US-09-414-189-1	Sequence 1, Appli
49	32	54.2	472	5	PCT-US94-09235-2	Sequence 2, Appli
50	32	54.2	531	2	US-08-724-974A-2	Sequence 2, Appli
51	32	54.2	568	4	US-09-252-991A-29237	Sequence 29237, A
52	32	54.2	640	2	US-08-671-978A-10	Sequence 10, Appl
53	32	54.2	775	3	US-09-305-640-2	Sequence 2, Appli
54	32	54.2	939	4	US-09-540-236-2399	Sequence 2399, Ap
55	32	54.2	1288	3	US-08-762-428A-6	Sequence 6, Appli
56	32	54.2	1306	3	US-08-999-774A-13	Sequence 13, Appl
57	32	54.2	1403	2	US-08-387-942C-3	Sequence 3, Appli
58	31	52.5	144	4	US-09-252-991A-17740	Sequence 17740, A
59	31	52.5	154	3	US-09-041-886-32	Sequence 32, Appl
60	31	52.5	207	4	US-08-811-519-25	Sequence 25, Appl
61	31	52.5	226	3	US-09-176-657-3	Sequence 3, Appli
62	31	52.5	226	4	US-09-421-299-3	Sequence 3, Appli
63	31	52.5	373	4	US-09-252-991A-21230	Sequence 21230, A
64	31	52.5	409	4	US-09-489-039A-14217	Sequence 14217, A
65	31	52.5	420	4	US-09-252-991A-19389	Sequence 19389, A
66	31	52.5	477	1	US-08-453-956-25	Sequence 25, Appl
67	31	52.5	477	1	US-08-086-631-25	Sequence 25, Appl
68	31	52.5	477	2	US-08-452-930-25	Sequence 25, Appl



69	31	52.5	477	5	PCT-US93-08174-25	Sequence 25, Appl
70	31	52.5	485	1	US-08-453-956-15	Sequence 15, Appl
71	31	52.5	485	1	US-08-086-631-15	Sequence 15, Appl
72	31	52.5	485	2	US-08-452-930-15	Sequence 15, Appl
73	31	52.5	485	5	PCT-US93-08174-15	Sequence 15, Appl
74	31	52.5	516	4	US-09-291-170A-1	Sequence 1, Appli
75	31	52.5	516	4	US-09-724-884-1	Sequence 1, Appli
76	31	52.5	531	4	US-09-134-001C-3574	Sequence 3574, Ap
77	31	52.5	545	4	US-09-604-957-4	Sequence 4, Appli
78	31	52.5	810	2	US-08-820-170A-34	Sequence 34, Appl
79	31	52.5	810	3	US-09-055-699-34	Sequence 34, Appl
80	31	52.5	810	3	US-09-273-565-34	Sequence 34, Appl
81	31	52.5	810	4	US-09-565-538-34	Sequence 34, Appl
82	31	52.5	810	4	US-09-661-468-34	Sequence 34, Appl
83	31	52.5	810	4	US-09-976-165-34	Sequence 34, Appl
84	31	52.5	900	4	US-09-328-352-8072	Sequence 8072, Ap
85	31	52.5	903	4	US-09-252-991A-28775	Sequence 28775, A
86	31	52.5	918	3	US-09-041-886-11	Sequence 11, Appl
87	31	52.5	1430	3	US-09-008-172-2	Sequence 2, Appli
88	31	52.5	1430	3	US-09-210-361-6	Sequence 6, Appli
89	31	52.5	1430	4	US-09-740-274-6	Sequence 6, Appli
90	31	52.5	1996	2	US-08-804-227C-9	Sequence 9, Appli
91	31	52.5	1996	2	US-08-804-198-3	Sequence 3, Appli
92	31	52.5	3460	4	US-09-334-220-1	Sequence 1, Appli
93	31	52.5	3461	4	US-09-334-220-2	Sequence 2, Appli
94	31	52.5	4545	2	US-08-804-227C-14	Sequence 14, Appl
95	31	52.5	4550	2	US-08-804-227C-8	Sequence 8, Appli
96	31	52.5	4550	2	US-08-804-198-2	Sequence 2, Appli
97	30	50.8	135	4	US-09-328-352-5042	Sequence 5042, Ap
98	30	50.8	136	4	US-08-826-134-15	Sequence 15, Appl
99	30	50.8	136	4	US-08-826-134-26	Sequence 26, Appl
100	30	50.8	162	4	US-09-732-210-527	Sequence 527, App
101	30	50.8	207	1	US-07-732-242C-5	Sequence 5, Appli
102	30	50.8	266	4	US-09-252-991A-32047	Sequence 32047, A
103	30	50.8	295	3	US-09-150-133-13	Sequence 13, Appl
104	30	50.8	295	3	US-09-150-141-13	Sequence 13, Appl
105	30	50.8	295	3	US-09-374-493-13	Sequence 13, Appl
106	30	50.8	295	3	US-09-374-824-13	Sequence 13, Appl
107	30	50.8	295	3	US-09-374-492-13	Sequence 13, Appl
108	30	50.8	295	4	US-09-785-343-13	Sequence 13, Appl
109	30	50.8	299	4	US-09-648-004-22	Sequence 22, Appl
110	30	50.8	345	4	US-09-252-991A-28350	Sequence 28350, A
111	30	50.8	347	4	US-09-489-039A-9485	Sequence 9485, Ap
112	30	50.8	365	4	US-09-721-870-36	Sequence 36, Appl
113	30	50.8	368	4	US-09-819-607-5	Sequence 5, Appli
114	30	50.8	413	4	US-09-252-991A-20866	Sequence 20866, A
115	30	50.8	413	4	US-09-491-577-72	Sequence 72, Appl
116	30	50.8	429	4	US-09-219-865B-12	Sequence 12, Appl
117	30	50.8	430	4	US-09-134-001C-4712	Sequence 4712, Ap
118	30	50.8	436	4	US-09-252-991A-28888	Sequence 28888, A
119	30	50.8	445	4	US-09-252-991A-18593	Sequence 18593, A
120	30	50.8	445	4	US-09-107-532A-5506	Sequence 5506, Ap
121	30	50.8	461	4	US-09-489-039A-13374	Sequence 13374, A
122	30	50.8	475	4	US-09-252-991A-18928	Sequence 18928, A
123	30	50.8	558	3	US-08-836-567-6	Sequence 6, Appli
124	30	50.8	558	4	US-09-606-304-6	Sequence 6, Appli
125	30	50.8	616	4	US-09-543-681A-4421	Sequence 4421, Ap

126	30	50.8	618	4	US-09-555-820A-14	Sequence 14, Appl
127	30	50.8	626	4	US-09-214-619-4	Sequence 4, Appli
128	30	50.8	626	4	US-09-555-820A-12	Sequence 12, Appl
129	30	50.8	688	4	US-09-252-991A-27096	Sequence 27096, A
130	30	50.8	689	3	US-09-106-194-2	Sequence 2, Appli
131	30	50.8	740	4	US-09-134-000C-6441	Sequence 6441, Ap
132	30	50.8	767	3	US-08-836-567-8	Sequence 8, Appli
133	30	50.8	767	4	US-09-606-304-8	Sequence 8, Appli
134	30	50.8	1142	4	US-09-106-075A-89	Sequence 89, Appl
135	30	50.8	1174	4	US-09-252-991A-29279	Sequence 29279, A
136	30	50.8	1333	4	US-09-347-878-20	Sequence 20, Appl
137	30	50.8	1381	4	US-08-826-134-4	Sequence 4, Appli
138	30	50.8	1384	4	US-08-826-134-2	Sequence 2, Appli
139	29.5	50.0	59	4	US-09-673-395A-423	Sequence 423, App
140	29.5	50.0	101	3	US-09-100-802-9	Sequence 9, Appli
141	29.5	50.0	102	4	US-09-149-476-366	Sequence 366, App
142	29	49.2	20	1	US-08-199-508-23	Sequence 23, Appl
143	29	49.2	26	1	US-07-791-930C-5	Sequence 5, Appli
144	29	49.2	26	1	US-08-173-515B-5	Sequence 5, Appli
145	29	49.2	69	4	US-08-311-731A-340	Sequence 340, App
146	29	49.2	111	4	US-09-621-976-6856	Sequence 6856, Ap
147	29	49.2	154	4	US-09-732-210-871	Sequence 871, App
148	29	49.2	161	4	US-09-252-991A-28429	Sequence 28429, A
149	29	49.2	175	4	US-09-479-313B-6	Sequence 6, Appli
150	29	49.2	175	4	US-09-479-313B-7	Sequence 7, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-313-288B-20

Query Match 94.9%; Score 56; DB 1; Length 1170;  
 Best Local Similarity 91.7%; Pred. No. 0.0068;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
 |||||:||||  
 Db 208 FQGVLSVRFVF 219

#### RESULT 2

US-09-252-991A-18826  
 ; Sequence 18826, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 18826  
 ; LENGTH: 175  
 ; TYPE: PRT  
 ; ORGANISM: Pseudomonas aeruginosa  
 US-09-252-991A-18826

Query Match 64.4%; Score 38; DB 4; Length 175;  
 Best Local Similarity 63.6%; Pred. No. 3.7;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLSVRFVF 12  
 ||||:|:|:  
 Db 41 QGVLSVRFVF 51

#### RESULT 3

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 20:55:10 ; Search time 33.1429 Seconds  
(without alignments)  
95.091 Million cell updates/sec

Title: US-10-030-735-21  
Perfect score: 59  
Sequence: 1 FQGVLSVRFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Published Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	

1	56	94.9	240	12	US-10-419-462-40	Sequence 40, Appl
2	56	94.9	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	56	94.9	831	12	US-09-939-853A-97	Sequence 97, Appl
4	56	94.9	831	12	US-09-939-853A-98	Sequence 98, Appl
5	56	94.9	1152	9	US-09-919-603-1	Sequence 1, Appli
6	56	94.9	1170	12	US-10-211-462-38	Sequence 38, Appl
7	56	94.9	1170	12	US-10-231-956A-482	Sequence 482, App
8	56	94.9	1170	12	US-10-419-462-38	Sequence 38, Appl
9	56	94.9	1170	14	US-10-020-141-12	Sequence 12, Appl
10	56	94.9	1170	14	US-10-017-721-2	Sequence 2, Appli
11	56	94.9	1170	14	US-10-021-660-114	Sequence 114, App
12	56	94.9	1170	14	US-10-008-093-2	Sequence 2, Appli
13	56	94.9	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	40	67.8	15	15	US-10-285-394-153	Sequence 153, App
15	38	64.4	16	9	US-09-822-682-6	Sequence 6, Appli
16	38	64.4	151	9	US-09-925-299-1251	Sequence 1251, Ap
17	38	64.4	151	10	US-09-925-299-1251	Sequence 1251, Ap
18	38	64.4	1168	9	US-09-919-603-2	Sequence 2, Appli
19	38	64.4	1172	9	US-09-919-770-4	Sequence 4, Appli
20	38	64.4	1172	9	US-09-822-682-2	Sequence 2, Appli
21	38	64.4	1172	10	US-09-866-050A-658	Sequence 658, App
22	38	64.4	1172	14	US-10-060-036-171	Sequence 171, App
23	38	64.4	1172	14	US-10-020-141-14	Sequence 14, Appl
24	38	64.4	1172	14	US-10-017-724-2	Sequence 2, Appli
25	38	64.4	1172	14	US-10-301-822-203	Sequence 203, App
26	38	64.4	1172	14	US-10-008-093-4	Sequence 4, Appli
27	38	64.4	1172	14	US-10-007-781-2	Sequence 2, Appli
28	38	64.4	1172	15	US-10-372-683-16	Sequence 16, Appl
29	38	64.4	1508	15	US-10-369-493-4104	Sequence 4104, Ap
30	36	61.0	215	14	US-10-029-386-33750	Sequence 33750, A
31	36	61.0	226	9	US-09-925-300-1217	Sequence 1217, Ap
32	36	61.0	390	14	US-10-420-845-6	Sequence 6, Appli
33	36	61.0	539	15	US-10-422-536-16	Sequence 16, Appl
34	36	61.0	559	14	US-10-284-499-47	Sequence 47, Appl
35	36	61.0	651	14	US-10-420-845-5	Sequence 5, Appli
36	36	61.0	1070	14	US-10-420-845-22	Sequence 22, Appl
37	36	61.0	1300	10	US-09-896-923-3	Sequence 3, Appli
38	36	61.0	1300	15	US-10-223-650-3	Sequence 3, Appli
39	36	61.0	1504	9	US-09-932-145-7	Sequence 7, Appli
40	36	61.0	1504	14	US-10-284-499-1	Sequence 1, Appli
41	36	61.0	1504	15	US-10-395-812-7	Sequence 7, Appli
42	36	61.0	1751	15	US-10-435-766-103	Sequence 103, App
43	35	59.3	51	12	US-10-424-599-198354	Sequence 198354,
44	35	59.3	1247	9	US-09-815-242-4886	Sequence 4886, Ap
45	35	59.3	1271	9	US-09-815-242-10717	Sequence 10717, A
46	35	59.3	1271	12	US-10-282-122A-42498	Sequence 42498, A
47	34	57.6	282	12	US-10-424-599-244115	Sequence 244115,
48	34	57.6	289	15	US-10-369-493-19908	Sequence 19908, A
49	34	57.6	329	12	US-10-424-599-177504	Sequence 177504,
50	34	57.6	375	12	US-10-424-599-226281	Sequence 226281,
51	34	57.6	428	15	US-10-421-654-28	Sequence 28, Appl
52	34	57.6	428	15	US-10-421-654-44	Sequence 44, Appl
53	34	57.6	435	15	US-10-369-493-3999	Sequence 3999, Ap
54	34	57.6	459	12	US-10-282-122A-67080	Sequence 67080, A
55	34	57.6	564	15	US-10-369-493-3903	Sequence 3903, Ap
56	34	57.6	1224	12	US-10-282-122A-70506	Sequence 70506, A

57	34	57.6	1240	12	US-10-282-122A-57585	Sequence 57585, A
58	33	55.9	50	12	US-10-424-599-193492	Sequence 193492,
59	33	55.9	75	12	US-10-424-599-165575	Sequence 165575,
60	33	55.9	83	12	US-10-424-599-150515	Sequence 150515,
61	33	55.9	211	12	US-10-335-977-7519	Sequence 7519, Ap
62	33	55.9	224	12	US-10-335-977-7521	Sequence 7521, Ap
63	33	55.9	263	15	US-10-369-493-4853	Sequence 4853, Ap
64	33	55.9	263	15	US-10-369-493-7612	Sequence 7612, Ap
65	33	55.9	265	12	US-10-282-122A-47647	Sequence 47647, A
66	33	55.9	277	9	US-09-738-626-4555	Sequence 4555, Ap
67	33	55.9	365	16	US-10-344-738-76	Sequence 76, Appl
68	33	55.9	367	12	US-10-282-122A-43731	Sequence 43731, A
69	33	55.9	370	9	US-09-895-913A-210	Sequence 210, App
70	33	55.9	406	12	US-10-335-977-7522	Sequence 7522, Ap
71	33	55.9	429	14	US-10-152-254-6	Sequence 6, Appli
72	33	55.9	459	15	US-10-369-493-18328	Sequence 18328, A
73	33	55.9	510	14	US-10-311-879-22	Sequence 22, Appl
74	33	55.9	520	12	US-10-424-599-209586	Sequence 209586,
75	33	55.9	525	12	US-10-425-114-55379	Sequence 55379, A
76	33	55.9	1213	9	US-09-815-242-5358	Sequence 5358, Ap
77	33	55.9	1217	9	US-09-815-242-12523	Sequence 12523, A
78	33	55.9	1217	12	US-10-282-122A-43796	Sequence 43796, A
79	33	55.9	1228	12	US-10-282-122A-71769	Sequence 71769, A
80	33	55.9	1829	15	US-10-435-766-13	Sequence 13, Appl
81	33	55.9	3500	14	US-10-153-219-2	Sequence 2, Appli
82	33	55.9	3537	14	US-10-153-219-15	Sequence 15, Appl
83	33	55.9	3586	15	US-10-334-143-77	Sequence 77, Appl
84	32	54.2	55	13	US-10-016-634A-119	Sequence 119, App
85	32	54.2	69	12	US-10-424-599-206194	Sequence 206194,
86	32	54.2	96	12	US-10-424-599-185817	Sequence 185817,
87	32	54.2	106	12	US-10-424-599-242530	Sequence 242530,
88	32	54.2	108	15	US-10-369-493-13199	Sequence 13199, A
89	32	54.2	140	10	US-09-796-753-136	Sequence 136, App
90	32	54.2	159	10	US-09-809-391-499	Sequence 499, App
91	32	54.2	159	10	US-09-882-171-499	Sequence 499, App
92	32	54.2	166	12	US-10-424-599-177709	Sequence 177709,
93	32	54.2	177	12	US-10-424-599-145722	Sequence 145722,
94	32	54.2	190	14	US-10-029-386-32211	Sequence 32211, A
95	32	54.2	191	12	US-10-072-012-26	Sequence 26, Appl
96	32	54.2	209	12	US-10-425-114-48779	Sequence 48779, A
97	32	54.2	215	12	US-10-072-012-28	Sequence 28, Appl
98	32	54.2	235	12	US-10-424-599-181083	Sequence 181083,
99	32	54.2	242	12	US-10-424-599-228787	Sequence 228787,
100	32	54.2	267	12	US-10-425-114-63233	Sequence 63233, A
101	32	54.2	278	9	US-09-778-927A-60	Sequence 60, Appl
102	32	54.2	278	12	US-10-262-839-150	Sequence 150, App
103	32	54.2	291	12	US-10-424-599-228788	Sequence 228788,
104	32	54.2	304	15	US-10-297-022-5	Sequence 5, Appli
105	32	54.2	306	14	US-10-128-714-3321	Sequence 3321, Ap
106	32	54.2	311	12	US-10-425-114-54870	Sequence 54870, A
107	32	54.2	320	15	US-10-369-493-17079	Sequence 17079, A
108	32	54.2	333	12	US-10-262-839-148	Sequence 148, App
109	32	54.2	343	9	US-09-925-301-1043	Sequence 1043, Ap
110	32	54.2	365	10	US-09-850-948-6	Sequence 6, Appli
111	32	54.2	365	10	US-09-964-821B-8	Sequence 8, Appli
112	32	54.2	365	14	US-10-109-533A-2	Sequence 2, Appli
113	32	54.2	365	14	US-10-251-385-118	Sequence 118, App

114	32	54.2	365	14	US-10-251-385-226	Sequence 226, App
115	32	54.2	365	14	US-10-273-575-6	Sequence 6, Appli
116	32	54.2	365	14	US-10-225-567A-283	Sequence 283, App
117	32	54.2	365	14	US-10-288-222A-12	Sequence 12, Appl
118	32	54.2	365	14	US-10-268-332-8	Sequence 8, Appli
119	32	54.2	425	14	US-10-147-026-10	Sequence 10, Appl
120	32	54.2	447	12	US-10-072-012-24	Sequence 24, Appl
121	32	54.2	449	10	US-09-798-723-4	Sequence 4, Appli
122	32	54.2	461	13	US-10-051-186-2	Sequence 2, Appli
123	32	54.2	461	14	US-10-225-567A-82	Sequence 82, Appl
124	32	54.2	461	14	US-10-021-660-121	Sequence 121, App
125	32	54.2	466	11	US-09-826-509-507	Sequence 507, App
126	32	54.2	466	14	US-10-225-567A-128	Sequence 128, App
127	32	54.2	466	14	US-10-241-220-78	Sequence 78, Appl
128	32	54.2	472	12	US-10-072-012-30	Sequence 30, Appl
129	32	54.2	472	12	US-10-211-462-73	Sequence 73, Appl
130	32	54.2	487	9	US-09-738-626-3647	Sequence 3647, Ap
131	32	54.2	492	14	US-10-152-254-7	Sequence 7, Appli
132	32	54.2	504	9	US-09-764-864-1173	Sequence 1173, Ap
133	32	54.2	508	14	US-10-128-714-8321	Sequence 8321, Ap
134	32	54.2	576	9	US-09-364-847-37	Sequence 37, Appl
135	32	54.2	710	12	US-10-282-122A-75098	Sequence 75098, A
136	32	54.2	712	9	US-09-364-847-49	Sequence 49, Appl
137	32	54.2	712	9	US-09-364-847-51	Sequence 51, Appl
138	32	54.2	775	10	US-09-798-723-2	Sequence 2, Appli
139	32	54.2	805	12	US-10-112-944-346	Sequence 346, App
140	32	54.2	1235	12	US-10-282-122A-60959	Sequence 60959, A
141	32	54.2	1252	12	US-10-282-122A-52163	Sequence 52163, A
142	32	54.2	1457	9	US-09-772-316-2	Sequence 2, Appli
143	32	54.2	1457	15	US-10-600-058-2	Sequence 2, Appli
144	32	54.2	1835	15	US-10-369-493-1489	Sequence 1489, Ap
145	31	52.5	40	10	US-09-974-879-219	Sequence 219, App
146	31	52.5	40	12	US-10-621-401-219	Sequence 219, App
147	31	52.5	41	10	US-09-305-736-219	Sequence 219, App
148	31	52.5	41	11	US-09-818-683-219	Sequence 219, App
149	31	52.5	50	9	US-09-971-980-64	Sequence 64, Appl
150	31	52.5	50	12	US-10-424-599-187405	Sequence 187405,

#### ALIGNMENTS

##### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 94.9%; Score 56; DB 12; Length 240;  
Best Local Similarity 91.7%; Pred. No. 0.0086;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
|||:||||  
Db 190 FQGVLSVRFVF 201

#### RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 94.9%; Score 56; DB 9; Length 466;  
Best Local Similarity 91.7%; Pred. No. 0.018;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
|||:||||  
Db 261 FQGVLSVRFVF 272

#### RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding  
Same



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:22:56 ; Search time 5.61039 Seconds  
(without alignments)  
111.372 Million cell updates/sec

Title: US-10-030-735-21  
Perfect score: 59  
Sequence: 1 FQGVLSVRFFV 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	56	94.9	1170	1	TSP1_BOVIN	Q28178 bos taurus
2	56	94.9	1170	1	TSP1_HUMAN	P07996 homo sapien
3	56	94.9	1170	1	TSP1_MOUSE	P35441 mus musculu
4	56	94.9	1173	1	TSP1_XENLA	P35448 xenopus lae
5	38	64.4	212	1	HIS5_NEIMA	Q9jvh3 neisseria m
6	38	64.4	1172	1	TSP2_HUMAN	P35442 homo sapien
7	38	64.4	1172	1	TSP2_MOUSE	Q03350 mus musculu
8	37	62.7	715	1	ZP2_CANFA	P47983 canis famil
9	36	61.0	1170	1	TSP2_BOVIN	Q95116 bos taurus
10	36	61.0	1178	1	TSP2_CHICK	P35440 gallus gall
11	36	61.0	1312	1	DPOL_PYRSD	Q51334 pyrococcus
12	36	61.0	1671	1	DPOL_PYRKO	P77933 pyrococcus
13	36	61.0	1699	1	DPOL_THEG8	Q9hh84 thermococcu
14	35	59.3	440	1	KRF1_SPVKA	P32216 swinepox vi
15	35	59.3	819	1	YCX1_CHLVU	P56370 chlorella v
16	34	57.6	355	1	GBA2_NEUCR	Q05424 neurospora
17	34	57.6	863	1	SIP1_YEAST	P32578 saccharomyc

18	34	57.6	1702	1	DPOL_THELI	P30317	thermococcu
19	33	55.9	467	1	FLII_BUCAI	P57178	buchnera ap
20	33	55.9	899	1	VP3_EHDV1	P27281	epizootic h
21	33	55.9	1134	1	YML7_YEAST	Q03735	saccharomyc
22	33	55.9	1205	1	PDS5_SCHPO	Q9hff5	schizosacch
23	33	55.9	1829	1	DPOL_THEST	O33845	thermococcu
24	33	55.9	1839	1	ALE3_AZOVI	Q44496	azotobacter
25	32	54.2	215	1	SHR3_SCHPO	Q9y876	schizosacch
26	32	54.2	252	1	YQGB_BACSU	P54485	bacillus su
27	32	54.2	295	1	LIPA_AERPE	Q9y9e3	aeropyrum p
28	32	54.2	333	1	CD1B_HUMAN	P29016	homo sapien
29	32	54.2	365	1	SPR1_HUMAN	Q15743	homo sapien
30	32	54.2	365	1	SPR1_MOUSE	Q8bfq3	mus musculu
31	32	54.2	369	1	YL05_SULSO	P95879	sulfolobus
32	32	54.2	385	1	YFD0_YEAST	P43567	saccharomyc
33	32	54.2	405	1	KRF1_VACCP	P29884	vaccinia vi
34	32	54.2	426	1	FOLC_BUCBP	Q89at2	buchnera ap
35	32	54.2	427	1	YIEM_ECOLI	P03818	escherichia
36	32	54.2	439	1	KRF1_VACCC	P21095	vaccinia vi
37	32	54.2	439	1	KRF1_VARV	P33801	variola vir
38	32	54.2	445	1	OXAA_THEMA	Q9x1h2	thermotoga
39	32	54.2	455	1	GIPR_RAT	P43219	rattus norv
40	32	54.2	461	1	CGRR_HUMAN	Q16602	homo sapien
41	32	54.2	462	1	GIPR_MESAU	P43218	mesocricetu
42	32	54.2	466	1	GIPR_HUMAN	P48546	homo sapien
43	32	54.2	607	1	LEPA_STAAM	Q99tr4	staphylococ
44	32	54.2	607	1	LEPA_STAAW	Q8nwa7	staphylococ
45	32	54.2	622	1	YA41_HUMAN	Q9upw0	homo sapien
46	32	54.2	775	1	ECEL_HUMAN	O95672	homo sapien
47	32	54.2	775	1	ECEL_MOUSE	Q9jmi0	mus musculu
48	32	54.2	899	1	VP3_EHDVA	P33474	epizootic h
49	32	54.2	901	1	VP3_BTV10	P12435	bluetongue
50	32	54.2	901	1	VP3_BTV11	Q65749	bluetongue
51	32	54.2	901	1	VP3_BTV13	Q65750	bluetongue
52	32	54.2	901	1	VP3_BTV17	P03539	bluetongue
53	32	54.2	901	1	VP3_BTV1A	P20608	bluetongue
54	32	54.2	901	1	VP3_BTV1S	P56582	bluetongue
55	32	54.2	1188	1	S3B1_SCHPO	Q10178	schizosacch
56	32	54.2	1288	1	OPLA_MOUSE	Q8k010	mus musculu
57	32	54.2	1288	1	OPLA_RAT	P97608	rattus norv
58	32	54.2	1304	1	S3B1_HUMAN	O75533	homo sapien
59	32	54.2	1304	1	S3B1_MOUSE	Q99nb9	mus musculu
60	32	54.2	1307	1	S3B1_XENLA	O57683	xenopus lae
61	32	54.2	1403	1	ALE1_AZOVI	Q44494	azotobacter
62	32	54.2	1835	1	DUR1_YEAST	P32528	saccharomyc
63	32	54.2	1995	1	YCX7_CHLRE	P36495	chlamydomon
64	31	52.5	146	1	JTB_MOUSE	O88824	mus musculu
65	31	52.5	196	1	BIM_MOUSE	O54918	mus musculu
66	31	52.5	196	1	BIM_RAT	O88498	rattus norv
67	31	52.5	197	1	PAT_ALCFA	P31668	alcaligenes
68	31	52.5	199	1	HIS5_HAEIN	P44340	haemophilus
69	31	52.5	211	1	PIMT_PSEAE	P45683	pseudomonas
70	31	52.5	268	1	RPNA_YEAST	P38886	saccharomyc
71	31	52.5	272	1	T2F1_HAEIN	P20588	haemophilus
72	31	52.5	273	1	DAPB_METTH	O26891	methanobact
73	31	52.5	275	1	UL34_HSV11	P10218	herpes simp
74	31	52.5	289	1	ATP6_PROMO	P21903	propionigen

75	31	52.5	312	1	OTCC_MYCCC	P59779	mycoplasma
76	31	52.5	313	1	MDH_AERPE	Q9yea1	aeropyrum p
77	31	52.5	317	1	RPC6_YEAST	P32910	saccharomyc
78	31	52.5	327	1	YRB2_YEAST	P40517	saccharomyc
79	31	52.5	357	1	MTDC_DROME	Q04448	d bifunctio
80	31	52.5	378	1	YIBH_ECOLI	P32107	escherichia
81	31	52.5	425	1	MURA_XYLFA	Q9pdg4	xylella fas
82	31	52.5	425	1	MURA_XYLFT	Q87dn8	xylella fas
83	31	52.5	455	1	GUDH_BACSU	P42238	bacillus su
84	31	52.5	474	1	NU4M_PARTE	P15581	paramecium
85	31	52.5	477	1	GLR_HUMAN	P47871	homo sapien
86	31	52.5	485	1	GLR_MOUSE	Q61606	mus musculu
87	31	52.5	485	1	GLR_RAT	P30082	rattus norv
88	31	52.5	511	1	MATK_ACCAL	Q9ghg9	acorus cala
89	31	52.5	526	1	MVIN_TREPA	O83529	treponema p
90	31	52.5	573	1	MLO4_ARATH	O23693	arabidopsis
91	31	52.5	607	1	LEPA_STAEP	Q8cp13	staphylococ
92	31	52.5	672	1	HYFB_ECOLI	P23482	escherichia
93	31	52.5	707	1	FLHA_TREPA	Q56338	treponema p
94	31	52.5	716	1	DVL3_HUMAN	Q92997	homo sapien
95	31	52.5	716	1	DVL3_MOUSE	Q61062	mus musculu
96	31	52.5	736	1	VP4_ROTTPC	P26193	porcine rot
97	31	52.5	810	1	NEL1_HUMAN	Q92832	homo sapien
98	31	52.5	810	1	NEL1_RAT	Q62919	rattus norv
99	31	52.5	820	1	AD29_HUMAN	Q9ukf5	homo sapien
100	31	52.5	871	1	UL47_HSVEB	P28929	equine herp
101	31	52.5	884	1	ANDR_EULFC	O97776	eulemur ful
102	31	52.5	894	1	OPLA_HUMAN	O14841	homo sapien
103	31	52.5	895	1	ANDR_MACFA	O97952	macaca fasc
104	31	52.5	895	1	ANDR_PAPHA	O97960	papio hamad
105	31	52.5	906	1	HELY_MYCTU	Q10701	mycobacteri
106	31	52.5	907	1	ANDR_CANFA	Q9tt90	canis famil
107	31	52.5	911	1	ANDR_PANTR	O97775	pan troglod
108	31	52.5	919	1	ANDR_HUMAN	P10275	homo sapien
109	31	52.5	920	1	HELY_MYCLE	Q9zbd8	mycobacteri
110	31	52.5	1041	1	RAG1_CHICK	P24271	gallus gall
111	31	52.5	1042	1	RAG1_RABIT	P34088	oryctolagus
112	31	52.5	1043	1	RAG1_HUMAN	P15918	homo sapien
113	31	52.5	1057	1	RAG1_BRARE	O13033	brachydanio
114	31	52.5	1147	1	DP3A_BORBU	O51526	borrelia bu
115	31	52.5	1177	1	Y307_MYCGE	P47549	mycoplasma
116	31	52.5	1462	1	GTFD_STRMU	P49331	streptococc
117	31	52.5	2027	1	DOC3_MOUSE	Q8ciq7	mus musculu
118	31	52.5	2030	1	DOC3_HUMAN	Q8izd9	homo sapien
119	31	52.5	3209	1	RELN_CHICK	O93574	gallus gall
120	31	52.5	3460	1	RELN_HUMAN	P78509	homo sapien
121	31	52.5	3461	1	RELN_MOUSE	Q60841	mus musculu
122	31	52.5	3462	1	RELN_RAT	P58751	rattus norv
123	30	50.8	162	1	RL10_HAEIN	P44350	haemophilus
124	30	50.8	163	1	RL10_PASMU	Q9ck89	pasteurella
125	30	50.8	204	1	UREG_BACSB	Q07403	bacillus sp
126	30	50.8	222	1	PLL2_MOUSE	P09586	mus musculu
127	30	50.8	226	1	ATP6_HELPJ	Q9zl15	helicobacte
128	30	50.8	226	1	ATP6_HELPY	P56085	helicobacte
129	30	50.8	259	1	TRMB_AZOSE	Q8g9c6	azoarcus sp
130	30	50.8	275	1	MURI_PSESM	Q888b8	pseudomonas
131	30	50.8	295	1	SUOT_MOUSE	P49891	mus musculu

132	30	50.8	297	1	NANA_SALTI	Q8z3f0	salmonella
133	30	50.8	297	1	NANA_SALTY	Q8z1q6	salmonella
134	30	50.8	297	1	Y4MH_RHISN	P55567	rhizobium s
135	30	50.8	338	1	PHND_ECOLI	P16682	escherichia
136	30	50.8	350	1	ENG_C_PASMU	Q9cmd1	pasteurella
137	30	50.8	352	1	IDI2_PYRAE	Q8zyf6	pyrobaculum
138	30	50.8	374	1	YF01_RHIME	Q92q49	rhizobium m
139	30	50.8	376	1	RECF_ANASP	Q8yrr9	anabaena sp
140	30	50.8	413	1	OR7A_DROME	Q9w3i5	drosophila
141	30	50.8	418	1	OR13_DROME	Q9vx10	drosophila
142	30	50.8	434	1	Y181_MYCPN	Q50292	mycoplasma
143	30	50.8	437	1	HEM1_ARCFU	O28304	archaeoglob
144	30	50.8	440	1	YA85_MYCPN	P75608	mycoplasma
145	30	50.8	466	1	FLII_BUCAP	Q8ka42	buchnera ap
146	30	50.8	492	1	CPAC_MOUSE	P56593	mus musculu
147	30	50.8	526	1	DPE2_MOUSE	O54956	mus musculu
148	30	50.8	549	1	Y447_MYCPN	P75130	mycoplasma
149	30	50.8	584	1	PARF_PARTE	P47244	paramecium
150	30	50.8	623	1	PGMP_ARATH	Q9scy0	arabidopsis

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.  
AC Q28178; Q28179;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Thrombospondin 1 precursor.  
GN THBS1 OR TSP1 OR TSP-1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Holstein; TISSUE=Tooth;  
RX MEDLINE=98173773; PubMed=9507054;  
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,  
RA Inoue H.;  
RT "cDNA cloning of bovine thrombospondin 1 and its expression in  
RT odontoblasts and predentin.";  
RL Biochim. Biophys. Acta 1382:17-22(1998).  
RN [2]  
RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.  
RC TISSUE=Aortic endothelium;  
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-  
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis  
CC and/or maintenance of dentin and dental pulp.

```

CC  -!- SUBUNIT: Homotrimer; disulfide-linked.
CC  -!- TISSUE SPECIFICITY: Odontoblasts.
CC  -!- SIMILARITY: Belongs to the thrombospondin family.
CC  -!- SIMILARITY: Contains 1 VWFC domain.
CC  -!- SIMILARITY: Contains 3 EGF-like domains.
CC  -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC  -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC  -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AB005287; BAA21115.1; -.
DR  EMBL; X87618; CAA60950.1; -.
DR  EMBL; X87619; CAA60951.1; -.
DR  PIR; S55501; S55501.
DR  GlycoSuiteDB; Q28178; -.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR006209; EGF_like.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000884; TSP1.
DR  InterPro; IPR008085; TSP_1.
DR  InterPro; IPR003367; tsp_3.
DR  InterPro; IPR008859; TSPC.
DR  InterPro; IPR003129; TSPN.
DR  InterPro; IPR001007; VWF_C.
DR  Pfam; PF00008; EGF; 2.
DR  Pfam; PF00090; tsp_1; 3.
DR  Pfam; PF02412; tsp_3; 13.
DR  Pfam; PF05735; TSPC; 1.
DR  Pfam; PF02210; TSPN; 1.
DR  Pfam; PF00093; vwc; 1.
DR  PRINTS; PR01705; TSP1REPEAT.
DR  SMART; SM00181; EGF; 3.
DR  SMART; SM00209; TSP1; 3.
DR  SMART; SM00210; TSPN; 1.
DR  SMART; SM00214; VWC; 1.
DR  PROSITE; PS00022; EGF_1; FALSE_NEG.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS50026; EGF_3; 2.
DR  PROSITE; PS50092; TSP1; 3.
DR  PROSITE; PS01208; VWFC_1; 1.
DR  PROSITE; PS50184; VWFC_2; 1.
KW  Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW  EGF-like domain; Signal.
FT  SIGNAL          1      18      BY SIMILARITY.
FT  CHAIN           19     1170     THROMBOSPONDIN 1.
FT  DOMAIN          19     232     HEPARIN-BINDING (POTENTIAL).
FT  DOMAIN          24     221     TSP N-TERMINAL.
FT  DOMAIN          316    373     VWFC.
FT  DOMAIN          379    429     TSP TYPE-1 1.
FT  DOMAIN          435    490     TSP TYPE-1 2.

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:33:49 ; Search time 31.2208 Seconds  
 (without alignments)  
 121.272 Million cell updates/sec

Title: US-10-030-735-21  
 Perfect score: 59  
 Sequence: 1 FQGVLSVRFVF 12  
 Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertibrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%				
No.	Score	Match	Length	DB	ID	Description
-----						

1	56	94.9	229	6	Q28194	Q28194 bos taurus
2	56	94.9	496	13	Q7SY84	Q7sy84 xenopus lae
3	56	94.9	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	56	94.9	1171	11	Q80YQ1	Q80yq1 mus musculu
5	38	64.4	145	5	Q93511	Q93511 caenorhabdi
6	38	64.4	146	17	Q976Q1	Q976q1 sulfolobus
7	38	64.4	278	10	Q94BI7	Q94bi7 celosia sp.
8	38	64.4	783	16	Q8UBI1	Q8ubi1 agrobacteri
9	38	64.4	1172	11	Q8CG21	Q8cg21 mus musculu
10	38	64.4	1172	11	Q7TMT3	Q7tmt3 mus musculu
11	38	64.4	1464	3	Q9C229	Q9c229 neurospora
12	37	62.7	394	6	Q8HZW7	Q8hzw7 ornithorhyn
13	37	62.7	659	16	Q8RCI0	Q8rci0 thermoanaer
14	37	62.7	689	16	Q8R808	Q8r808 thermoanaer
15	37	62.7	727	17	Q8PRY3	Q8pry3 methanosarc
16	37	62.7	733	17	Q8TLX6	Q8tlx6 methanosarc
17	37	62.7	750	11	Q8BVV4	Q8bvv4 mus musculu
18	37	62.7	957	13	Q8QG19	Q8qg19 menura nova
19	37	62.7	1711	9	Q8LTK2	Q8ltk2 lactococcus
20	37	62.7	1713	9	Q94MA1	Q94ma1 lactococcus
21	36	61.0	210	2	Q9ZEN6	Q9zen6 wolinnella s
22	36	61.0	219	16	Q8ENF8	Q8enf8 oceanobacil
23	36	61.0	325	13	Q90XK6	Q90xk6 amia calva
24	36	61.0	394	16	Q87M68	Q87m68 vibrio para
25	36	61.0	405	17	Q8U2Q0	Q8u2q0 pyrococcus
26	36	61.0	595	4	Q9UEU4	Q9ueu4 homo sapien
27	36	61.0	956	13	Q8QG47	Q8qg47 aegithalos
28	36	61.0	1142	3	Q9UT41	Q9ut41 schizosacch
29	36	61.0	1354	11	Q9EPW8	Q9epw8 mus musculu
30	36	61.0	1480	11	Q80TM9	Q80tm9 mus musculu
31	36	61.0	1504	4	Q9UES6	Q9ues6 homo sapien
32	36	61.0	1504	4	Q7Z2X6	Q7z2x6 homo sapien
33	36	61.0	1528	4	Q9Y2I1	Q9y2i1 homo sapien
34	35.5	60.2	1554	5	Q7YY16	Q7yy16 cryptospori
35	35	59.3	82	10	Q9ZUI2	Q9zui2 arabidopsis
36	35	59.3	217	16	Q8RC20	Q8rc20 thermoanaer
37	35	59.3	248	13	Q7ZWN5	Q7zwn5 xenopus lae
38	35	59.3	323	13	Q90XK7	Q90xk7 lepisosteus
39	35	59.3	460	10	Q94DF6	Q94df6 oryza sativ
40	35	59.3	465	16	Q8ZCV8	Q8zcv8 yersinia pe
41	35	59.3	593	2	Q8GH66	Q8gh66 mycobacteri
42	35	59.3	610	16	Q822D4	Q822d4 chlamydophi
43	35	59.3	627	16	Q87GG1	Q87gg1 vibrio para
44	35	59.3	657	5	Q8MNS5	Q8mns5 caenorhabdi
45	35	59.3	683	5	Q8MNS4	Q8mns4 caenorhabdi
46	35	59.3	849	5	Q9V780	Q9v780 drosophila
47	35	59.3	951	13	Q8QFY3	Q8qfy3 vanga curvi
48	35	59.3	952	13	Q8QFY5	Q8qfy5 troglodytes
49	35	59.3	952	13	Q8QFY9	Q8qfy9 thamnophilu
50	35	59.3	953	13	Q8QG14	Q8qg14 nectarinia
51	35	59.3	955	13	Q8QG08	Q8qg08 pardalotus
52	35	59.3	956	13	Q8QFY7	Q8qfy7 toxorhamphu
53	35	59.3	956	13	Q9W6P9	Q9w6p9 tyrannus ty
54	35	59.3	957	13	Q8QFZ5	Q8qfz5 regulus cal
55	35	59.3	957	13	Q8QG23	Q8qg23 lanius ludo
56	35	59.3	957	13	Q8QFZ4	Q8qfz4 rupicola ru
57	35	59.3	957	13	Q8QG38	Q8qg38 cinclus cin

58	35	59.3	957	13	Q8QG12	Q8qg12	oriolus lar
59	35	59.3	957	13	Q8QG37	Q8qg37	cisticola a
60	35	59.3	957	13	Q8QG10	Q8qg10	pachycephal
61	35	59.3	957	13	Q8QFZ1	Q8qfz1	sturnus vul
62	35	59.3	957	13	Q8QG04	Q8qg04	picathartes
63	35	59.3	957	13	Q8QG46	Q8qg46	aegithina t
64	35	59.3	957	13	Q8QG15	Q8qg15	muscipapa s
65	35	59.3	957	13	Q8QG26	Q8qg26	hirundo pyr
66	35	59.3	957	13	Q8QG28	Q8qg28	furnarius r
67	35	59.3	957	13	Q8QG24	Q8qg24	irena cyano
68	35	59.3	957	13	Q8QG21	Q8qg21	melanochari
69	35	59.3	957	13	Q8QG27	Q8qg27	garrulax mi
70	35	59.3	957	13	Q8QG43	Q8qg43	artamus leu
71	35	59.3	957	13	Q8QFZ0	Q8qfz0	sylvia nana
72	35	59.3	957	13	Q8QG13	Q8qg13	oedistoma i
73	35	59.3	957	13	Q8QG30	Q8qg30	formicarius
74	35	59.3	957	13	Q8QG18	Q8qg18	mimus patag
75	35	59.3	957	13	Q8QG03	Q8qg03	pipra coron
76	35	59.3	957	13	Q8QG11	Q8qg11	orthonyx sp
77	35	59.3	957	13	Q8QFY1	Q8qfy1	zosterops s
78	35	59.3	957	13	Q8QG17	Q8qg17	monarcha ax
79	35	59.3	957	13	Q8QG45	Q8qg45	alauda arve
80	35	59.3	957	13	Q8QG35	Q8qg35	coracina li
81	35	59.3	961	13	Q8QG07	Q8qg07	parula amer
82	35	59.3	961	13	Q8QFY8	Q8qfy8	thraupis cy
83	35	59.3	961	13	Q8QG41	Q8qg41	cardinalis
84	35	59.3	961	13	Q8QG31	Q8qg31	emberiza sc
85	35	59.3	961	13	Q9W6P8	Q9w6p8	passer mont
86	35	59.3	995	5	Q8SZP4	Q8szp4	drosophila
87	35	59.3	995	5	Q9V7I8	Q9v7i8	drosophila
88	35	59.3	1249	16	Q88U41	Q88u41	lactobacill
89	35	59.3	1264	16	Q836J8	Q836j8	enterococcu
90	34	57.6	151	16	Q7UDW7	Q7udw7	rhodopirell
91	34	57.6	204	5	Q8ST26	Q8st26	dictyosteli
92	34	57.6	207	5	O96097	O96097	antheraea y
93	34	57.6	207	5	O96096	O96096	antheraea p
94	34	57.6	212	16	Q8YTA4	Q8yta4	anabaena sp
95	34	57.6	214	16	Q82ZE3	Q82ze3	enterococcu
96	34	57.6	223	10	Q9LMB5	Q9lmb5	arabidopsis
97	34	57.6	243	11	Q99NN3	Q99nn3	dipodomys h
98	34	57.6	258	6	Q9BEX1	Q9bex1	nycteris th
99	34	57.6	262	6	Q8WN46	Q8wn46	nycteris th
100	34	57.6	265	16	P70972	P70972	bacillus su
101	34	57.6	279	10	Q94BF8	Q94bf8	gomphrena s
102	34	57.6	280	10	Q94BB4	Q94bb4	psilotrichu
103	34	57.6	280	10	Q94BE9	Q94be9	iresine her
104	34	57.6	282	10	Q94BK9	Q94bk9	alternanthe
105	34	57.6	288	10	Q94BE4	Q94be4	limeum afri
106	34	57.6	289	10	Q94BF9	Q94bf9	glischrotha
107	34	57.6	322	13	Q90XJ7	Q90xj7	typhlonecte
108	34	57.6	323	16	Q7WHU3	Q7whu3	bordetella
109	34	57.6	323	16	Q7W9Y4	Q7w9y4	bordetella
110	34	57.6	324	13	Q90XJ6	Q90xj6	latimeria m
111	34	57.6	348	16	O67217	O67217	aquifex aeo
112	34	57.6	359	6	Q8WN53	Q8wn53	nycteris gr
113	34	57.6	363	16	Q929N4	Q929n4	listeria in
114	34	57.6	364	16	Q8Y5D0	Q8y5d0	listeria mo



115	34	57.6	374	6	Q9TT59	Q9tt59 tadarida br
116	34	57.6	394	6	Q8HZX0	Q8hzx0 tachyglossu
117	34	57.6	394	6	Q7YSB6	Q7ysb6 zaglossus b
118	34	57.6	412	10	Q7XXL5	Q7xxl5 oryza sativ
119	34	57.6	412	16	Q8DUJ9	Q8duj9 streptococc
120	34	57.6	417	16	Q9KFK7	Q9kfk7 bacillus ha
121	34	57.6	456	8	Q9BA99	Q9ba99 pentachondr
122	34	57.6	459	16	Q9CMS7	Q9cms7 pasteurella
123	34	57.6	466	8	Q9BAA3	Q9baa3 pentachondr
124	34	57.6	480	8	Q9BA97	Q9ba97 pentachondr
125	34	57.6	485	8	O78396	O78396 brachyloma
126	34	57.6	489	8	Q9BA98	Q9ba98 pentachondr
127	34	57.6	492	13	Q9PUK1	Q9puk1 hoplobatrac
128	34	57.6	493	8	Q9BAA5	Q9baa5 pentachondr
129	34	57.6	587	5	Q86JT5	Q86jt5 dictyosteli
130	34	57.6	595	5	Q86NR6	Q86nr6 drosophila
131	34	57.6	721	5	Q9VTH0	Q9vth0 drosophila
132	34	57.6	787	16	Q8K6W1	Q8k6w1 streptococc
133	34	57.6	893	16	Q9S258	Q9s258 streptomyc
134	34	57.6	956	13	Q8QFZ7	Q8qfz7 ptilonorhyn
135	34	57.6	957	13	Q8QG32	Q8qg32 dicrurus ad
136	34	57.6	957	13	Q8QG09	Q8qg09 paradisaea
137	34	57.6	957	13	Q8QG02	Q8qg02 pitta guaja
138	34	57.6	957	13	Q8QG20	Q8qg20 meliphaga a
139	34	57.6	957	13	Q8QFZ3	Q8qfz3 sitta pygma
140	34	57.6	957	13	Q8QFZ8	Q8qfz8 psarisomus
141	34	57.6	957	13	Q8QG05	Q8qg05 philepitta
142	34	57.6	957	13	Q8QG36	Q8qg36 climacteris
143	34	57.6	957	13	Q8QG48	Q8qg48 acanthisitt
144	34	57.6	957	13	Q8QFZ2	Q8qfz2 smithornis
145	34	57.6	957	13	Q8QFY6	Q8qfy6 tregellasia
146	34	57.6	957	13	Q8QG33	Q8qg33 dicaeum mel
147	34	57.6	957	13	Q8QG39	Q8qg39 chloropsis
148	34	57.6	957	13	Q8QG40	Q8qg40 certhia fam
149	34	57.6	957	13	Q8QG22	Q8qg22 malurus mel
150	34	57.6	957	13	Q8QG00	Q8qg00 pomatostomu

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.

AC Q28194;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Thrombospondin-1 (Fragment).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 94.9%; Score 56; DB 6; Length 229;  
 Best Local Similarity 91.7%; Pred. No. 0.0058;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
 |||||:|||||  
 Db 190 FQGVLNVRVFVF 201

## RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipioidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC054970; AAH54970.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 496 AA; 54843 MW; E4FD2F07CB7EF51B CRC64;

Query Match 94.9%; Score 56; DB 13; Length 496;  
 Best Local Similarity 91.7%; Pred. No. 0.013;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
 |||||:|||||  
 Db 214 FQGVLSVRFVF 225

# RESULT 3

## Q8CGB2

ID Q8CGB2 PRELIMINARY; PRT; 1171 AA.  
 AC Q8CGB2;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Thrombospondin 1.  
 GN THBS1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CZECH II;  
 RA Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC042422; AAH42422.1; -.  
 DR MGD; MGI:98737; Thbs1.  
 DR GO; GO:0005615; C:extracellular space; IDA.  
 DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.  
 DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008859; TSPC.  
 DR InterPro; IPR003129; TSPN.  
 DR InterPro; IPR008085; TSP\_1.  
 DR InterPro; IPR003367; tsp\_3.  
 DR InterPro; IPR001007; VWF\_C.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF05735; TSPC; 1.  
 DR Pfam; PF02210; TSPN; 1.  
 DR Pfam; PF00090; tsp\_1; 3.  
 DR Pfam; PF02412; tsp\_3; 13.  
 DR Pfam; PF00093; vwc; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00181; EGF; 3.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00209; TSP1; 3.  
 DR SMART; SM00210; TSPN; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS01208; VWFC\_1; 1.  
 DR PROSITE; PS50184; VWFC\_2; 1.  
 SQ SEQUENCE 1171 AA; 129717 MW; 431C978460AB2C41 CRC64;

Query Match 94.9%; Score 56; DB 11; Length 1171;  
 Best Local Similarity 91.7%; Pred. No. 0.031;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSQSVRFVF 12  
 |||||:|||||  
 Db 208 FQGVLQNVRFVF 219

#### RESULT 4

##### Q80YQ1

ID Q80YQ1 PRELIMINARY; PRT; 1171 AA.  
 AC Q80YQ1;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Thrombospondin 1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC050917; AAH50917.1; -.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR003129; TSPN.  
 DR InterPro; IPR008085; TSP\_1.  
 DR InterPro; IPR003367; tsp\_3.  
 DR InterPro; IPR001007; VWF\_C.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF02210; TSPN; 1.  
 DR Pfam; PF00090; tsp\_1; 3.  
 DR Pfam; PF02412; tsp\_3; 8.  
 DR Pfam; PF00093; vwc; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00181; EGF; 3.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00209; TSP1; 3.  
 DR SMART; SM00210; TSPN; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS01208; VWFC\_1; 1.  
 DR PROSITE; PS50184; VWFC\_2; 1.  
 SQ SEQUENCE 1171 AA; 129689 MW; 12E077B50C64E2D3 CRC64;

Query Match 94.9%; Score 56; DB 11; Length 1171;  
 Best Local Similarity 91.7%; Pred. No. 0.031;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
 |||||:|||||  
 Db 208 FQGVLSVRFVF 219

# RESULT 5

Q93511

ID Q93511 PRELIMINARY; PRT; 145 AA.  
 AC Q93511;  
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE F15D4.3 protein.  
 GN F15D4.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Smye R.;  
 RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).

DR EMBL; Z80344; CAB02486.2; -.  
DR PIR; B88349; B88349.  
DR PIR; T20985; T20985.  
DR WormPep; F15D4.3; CE23658.  
SQ SEQUENCE 145 AA; 15766 MW; 285D67A56AED28DF CRC64;

Query Match 64.4%; Score 38; DB 5; Length 145;  
Best Local Similarity 50.0%; Pred. No. 14;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
| | | | : | : : |  
Db 73 FMGVAQGLRYIF 84

#### RESULT 6

Q976Q1  
ID Q976Q1 PRELIMINARY; PRT; 146 AA.  
AC Q976Q1;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein ST0139.  
GN ST0139.  
OS Sulfolobus tokodaii.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.  
OX NCBI\_TaxID=111955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JCM 10545 / 7;  
RX MEDLINE=21456156; PubMed=11572479;  
RA Kawarabayasi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,  
RA Sekine M., Baba S.-I., Ankai A., Kosugi H., Hosoyama A., Fukui S.,  
RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,  
RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kushida N., Oguchi A.,  
RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,  
RA Oshima T., Kikuchi H.;  
RT "Complete genome sequence of an aerobic thermoacidophilic  
RT Crenarchaeon, Sulfolobus tokodaii strain7.";  
RL DNA Res. 8:123-140(2001).  
DR EMBL; AP000981; BAB65095.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 146 AA; 16564 MW; F738A8B7151CE598 CRC64;

Query Match 64.4%; Score 38; DB 17; Length 146;  
Best Local Similarity 50.0%; Pred. No. 14;  
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
| | | | : | : : |  
Db 107 FLGVLEKVKFIY 118

#### RESULT 7

Q94BI7  
ID Q94BI7 PRELIMINARY; PRT; 278 AA.

AC Q94BI7;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Maturase K (Intron maturase) (Fragment).  
 GN MATK.  
 OS Celosia sp. Cuenoud s.n.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Celosia.  
 OX NCBI\_TaxID=169523;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Cuenoud P., Savolainen V., Chatrou L.W., Powell M., Grayer R.J.,  
 RA Chase M.W.;  
 RT "Molecular phylogenetics of the Caryophyllales based on 18S rDNA,  
 RT rbcL, atpB, and matK sequences."  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II  
 CC INTRONS (BY SIMILARITY).  
 CC -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,  
 CC AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY  
 CC MITOCHONDRIAL INTRONS.  
 DR EMBL; AY042565; AAK94551.1; -.  
 DR GO; GO:0009507; C:chloroplast; IEA.  
 DR GO; GO:0006397; P:mRNA processing; IEA.  
 DR GO; GO:0008380; P:RNA splicing; IEA.  
 DR InterPro; IPR000442; Intron\_maturase2.  
 DR InterPro; IPR002866; MatK\_N.  
 DR Pfam; PF01348; Intron\_maturase2; 1.  
 DR Pfam; PF01824; MatK\_N; 1.  
 KW mRNA processing; Chloroplast.  
 FT NON\_TER 1 1  
 FT NON\_TER 278 278  
 SQ SEQUENCE 278 AA; 33022 MW; B98513EAC40E3074 CRC64;

Query Match 64.4%; Score 38; DB 10; Length 278;  
 Best Local Similarity 72.7%; Pred. No. 28;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLSVRFV 11  
 | | | | |  
 Db 186 FMGFLSSVRFV 196

#### RESULT 8

##### Q8UBI1

ID Q8UBI1 PRELIMINARY; PRT; 783 AA.  
 AC Q8UBI1;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Two component response regulator.  
 GN ATU3035 OR AGR\_L\_3540.  
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.

OX NCBI\_TaxID=176299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608550; PubMed=11743193;  
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,  
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,  
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,  
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,  
 RA Kuttyavin T., Levy R., Li M.-J., McClelland E., Palmieri A.,  
 RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,  
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,  
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,  
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,  
 RA Nester E.W.;  
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*  
 RT C58.";  
 RL Science 294:2317-2323(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21608551; PubMed=11743194;  
 RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,  
 RA Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,  
 RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,  
 RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,  
 RA Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,  
 RA Cielo C., Slater S.;  
 RT "Genome sequence of the plant pathogen and biotechnology agent  
 RT *Agrobacterium tumefaciens* C58.";  
 RL Science 294:2323-2328(2001).  
 DR EMBL; AE009233; AAL43851.1; ALT\_INIT.  
 DR EMBL; AE008380; AAK90347.1; -.  
 DR PIR; A98353; A98353.  
 DR PIR; AE2929; AE2929.  
 DR GO; GO:0005622; C:intracellular; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro; IPR000792; HTH\_LuxR.  
 DR Pfam; PF00196; GerE; 1.  
 DR PRINTS; PR00038; HTHLUXR.  
 DR ProDom; PD000307; HTH\_LuxR; 1.  
 DR SMART; SM00421; HTH\_LUXR; 1.  
 DR PROSITE; PS00622; HTH\_LUXR\_FAMILY; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 783 AA; 88637 MW; 0C3F5862C135EC94 CRC64;

Query Match 64.4%; Score 38; DB 16; Length 783;  
 Best Local Similarity 77.8%; Pred. No. 81;  
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GVLQSVRFV 11  
 |:|:|||||  
 Db 78 GILES SVRFV 86

RESULT 9  
 Q8CG21  
 ID Q8CG21 PRELIMINARY; PRT; 1172 AA.



AC Q8CG21;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Thbs2.  
 GN THBS2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129S6/SvEvTac;  
 RA Brathwaite M., Waeltz P., Qian Y., Dudekula D., Schlessinger D.,  
 RA Nagaraja R.;  
 RT "Genomic Sequence Analysis in the Mouse t-complex Region.";  
 RL Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF549256; AA016244.1; -.  
 DR GO; GO:0005509; F:calcium ion binding; IEA.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR001881; EGF\_Ca.  
 DR InterPro; IPR006209; EGF\_like.  
 DR InterPro; IPR006210; IEGF.  
 DR InterPro; IPR000884; TSP1.  
 DR InterPro; IPR008859; TSPC.  
 DR InterPro; IPR003129; TSPN.  
 DR InterPro; IPR008085; TSP\_1.  
 DR InterPro; IPR003367; tsp\_3.  
 DR InterPro; IPR001007; VWF\_C.  
 DR Pfam; PF00008; EGF; 2.  
 DR Pfam; PF05735; TSPC; 1.  
 DR Pfam; PF02210; TSPN; 1.  
 DR Pfam; PF00090; tsp\_1; 3.  
 DR Pfam; PF02412; tsp\_3; 13.  
 DR Pfam; PF00093; vwc; 1.  
 DR PRINTS; PR01705; TSP1REPEAT.  
 DR SMART; SM00181; EGF; 3.  
 DR SMART; SM00179; EGF\_CA; 2.  
 DR SMART; SM00209; TSP1; 3.  
 DR SMART; SM00210; TSPN; 1.  
 DR SMART; SM00214; VWC; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS50092; TSP1; 3.  
 DR PROSITE; PS01208; VWFC\_1; 1.  
 DR PROSITE; PS50184; VWFC\_2; 1.  
 SQ SEQUENCE 1172 AA; 129881 MW; 020ACD7EB5137B25 CRC64;

Query Match 64.4%; Score 38; DB 11; Length 1172;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLSVRFVF 12  
 |:|:|:| ||  
 Db 202 FRGLLQNVHLVF 213

RESULT 10

Q7TMT3

ID Q7TMT3 PRELIMINARY; PRT; 1172 AA.  
 AC Q7TMT3;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Thrombospondin 2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/He; TISSUE=Mesenchymal Stem Cell;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C3H/He; TISSUE=Mesenchymal Stem Cell;  
 RA Strausberg R.;  
 RL Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; BC053702; AAH53702.1; -.  
 SQ SEQUENCE 1172 AA; 129942 MW; B620EF2B2F5DF0D5 CRC64;

Query Match 64.4%; Score 38; DB 11; Length 1172;  
 Best Local Similarity 58.3%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
 |:|:|:|:| ||  
 Db 202 FRGLLQNVHLVF 213

RESULT 11

Q9C229

ID Q9C229 PRELIMINARY; PRT; 1464 AA.  
 AC Q9C229;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Related to ATP-binding cassette transporter protein YOR1.  
 GN B7A16.190.  
 OS Neurospora crassa.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.  
 OX NCBI\_TaxID=5141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Schulte U., Aign V., Hoheisel J., Brandt P., Fartmann B., Holland R.,  
 RA Nyakatuta G., Mewes H.W., Mannhaupt G.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA German Neurospora genome project;  
 RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.  
 DR EMBL; AL513445; CAC28731.2; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
 DR GO; GO:0000166; F:nucleotide binding; IEA.  
 DR GO; GO:0006810; P:transport; IEA.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR001140; ABC\_TM\_transpt.  
 DR InterPro; IPR003439; ABC\_transporter.  
 DR InterPro; IPR001092; HLH\_basic.  
 DR Pfam; PF00664; ABC\_membrane; 2.  
 DR Pfam; PF00005; ABC\_tran; 2.  
 DR ProDom; PD000006; ABC\_transporter; 2.  
 DR SMART; SM00382; AAA; 2.  
 DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 2.  
 DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 2.  
 DR PROSITE; PS00038; HLH\_1; 1.  
 KW ATP-binding; Transport.  
 SQ SEQUENCE 1464 AA; 162131 MW; 53D5470B17734385 CRC64;

Query Match 64.4%; Score 38; DB 3; Length 1464;  
 Best Local Similarity 80.0%; Pred. No. 1.5e+02;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 QGVLQSVRFV 11  
 | :|||||  
 Db 407 QEILQSVRFV 416

# RESULT 12

Q8HZW7

ID Q8HZW7 PRELIMINARY; PRT; 394 AA.

AC Q8HZW7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Recombination activating gene-1 (Fragment).

GN RAG-1.

OS Ornithorhynchus anatinus (Duckbill platypus).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Monotremata; Ornithorhynchidae; Ornithorhynchus.  
 OX NCBI\_TaxID=9258;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Baker M.L., Miller R.D.;  
 RT "Characterization of RAG1 genes from metatherians and monotremes.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF303974; AAL09342.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 394 394  
 SQ SEQUENCE 394 AA; 45838 MW; A63AF68169C42A4D CRC64;

Query Match 62.7%; Score 37; DB 6; Length 394;  
 Best Local Similarity 60.0%; Pred. No. 63;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|:| ||:|  
 Db 79 GILRSFRFIF 88

# RESULT 13

## Q8RCI0

ID Q8RCI0 PRELIMINARY; PRT; 659 AA.  
 AC Q8RCI0;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Methyl-accepting chemotaxis protein.  
 GN TAR OR TTE0448.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome.";  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL; AE013016; AAM23732.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0004871; F:signal transducer activity; IEA.  
 DR GO; GO:0006935; P:chemotaxis; IEA.  
 DR GO; GO:0007165; P:signal transduction; IEA.  
 DR InterPro; IPR004010; Cache.  
 DR InterPro; IPR004089; Chmtaxis\_transd.  
 DR InterPro; IPR003660; HAMP.  
 DR Pfam; PF02743; Cache; 1.  
 DR Pfam; PF00672; HAMP; 1.  
 DR Pfam; PF00015; MCPsignal; 1.  
 DR SMART; SM00304; HAMP; 1.

DR SMART; SM00283; MA; 1.  
DR PROSITE; PS50111; CHEMOTAXIS\_TRANSDUC\_2; 1.  
DR PROSITE; PS50885; HAMP; 1.  
KW Complete proteome.  
SQ SEQUENCE 659 AA; 72320 MW; 63A2826249EDA03E CRC64;

Query Match 62.7%; Score 37; DB 16; Length 659;  
Best Local Similarity 45.5%; Pred. No. 1.1e+02;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLSVRFV 11  
|:|:|:|:|:  
Db 569 FEGILNAVKFI 579

#### RESULT 14

Q8R808

ID Q8R808 PRELIMINARY; PRT; 689 AA.  
AC Q8R808;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Methyl-accepting chemotaxis protein.  
GN TAR9 OR TTE2231.  
OS Thermoanaerobacter tengcongensis.  
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
OX NCBI\_TaxID=119072;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MB4 / JCM 11007;  
RX MEDLINE=21992816; PubMed=11997336;  
RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
RT "A complete sequence of T. tengcongensis genome.";  
RL Genome Res. 12:689-700(2002).  
DR EMBL; AE013167; AAM25381.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0004871; F:signal transducer activity; IEA.  
DR GO; GO:0006935; P:chemotaxis; IEA.  
DR GO; GO:0007165; P:signal transduction; IEA.  
DR InterPro; IPR004010; Cache.  
DR InterPro; IPR004089; Chmtaxis\_transd.  
DR InterPro; IPR003660; HAMP.  
DR Pfam; PF02743; Cache; 1.  
DR Pfam; PF00672; HAMP; 1.  
DR Pfam; PF00015; MCPsignal; 1.  
DR SMART; SM00304; HAMP; 1.  
DR SMART; SM00283; MA; 1.  
DR PROSITE; PS50111; CHEMOTAXIS\_TRANSDUC\_2; 1.  
DR PROSITE; PS50885; HAMP; 1.  
KW Complete proteome.  
SQ SEQUENCE 689 AA; 75721 MW; 25C956672E8BDAC5 CRC64;

Query Match 62.7%; Score 37; DB 16; Length 689;  
Best Local Similarity 45.5%; Pred. No. 1.1e+02;

Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 11  
|:|:| :|:|:  
Db 599 FEGILNAVKFI 609

RESULT 15

Q8PRY3

ID Q8PRY3 PRELIMINARY; PRT; 727 AA.  
AC Q8PRY3;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative DNA or RNA helicase of superfamily II.  
GN MM3305.  
OS Methanosarcina mazei (Methanosarcina frisia).  
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
OX NCBI\_TaxID=2209;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Goe1 / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;  
RX MEDLINE=22120827; PubMed=12125824;  
RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,  
RA Martinez-Arias R., Henne A., Wiezer A., Baeumer S., Jacobi C.,  
RA Brueggemann H., Lienard T., Christmann A., Boemecke M., Steckel S.,  
RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,  
RA Fritz H.-J., Gottschalk G.;  
RT "The genome of Methanosarcina mazei: evidence for lateral gene  
RT transfer between Bacteria and Archaea."  
RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).  
DR EMBL; AE013591; AAM33001.1; -.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004386; F:helicase activity; IEA.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR InterPro; IPR001650; Helicase\_C.  
DR Pfam; PF00271; helicase\_C; 1.  
KW Helicase; Complete proteome; Hypothetical protein.  
SQ SEQUENCE 727 AA; 83319 MW; A3CA3123E3282060 CRC64;

Query Match 62.7%; Score 37; DB 17; Length 727;  
Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 11  
||||: :|:|  
Db 186 FQGVVGNVKFV 196

RESULT 16

Q8TLX6

ID Q8TLX6 PRELIMINARY; PRT; 733 AA.  
AC Q8TLX6;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein MA2897.  
 GN MA2897.  
 OS Methanosarcina acetivorans.  
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
 RX MEDLINE=21929760; PubMed=11932238;  
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,  
 RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,  
 RA Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,  
 RA Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,  
 RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RT "The genome of Methanosarcina acetivorans reveals extensive metabolic  
 RT and physiological diversity."  
 RL Genome Res. 12:532-542(2002).  
 DR EMBL; AE010991; AAM06273.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0008026; F:ATP dependent helicase activity; IEA.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR InterPro; IPR001410; DEAD.  
 DR InterPro; IPR001650; Helicase\_C.  
 DR Pfam; PF00271; helicase\_C; 1.  
 DR SMART; SM00487; DEXDc; 1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 733 AA; 83531 MW; A2B7B4EE81454FA3 CRC64;  
  
 Query Match 62.7%; Score 37; DB 17; Length 733;  
 Best Local Similarity 63.6%; Pred. No. 1.2e+02;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
 Qy 1 FQGVLSVRFV 11  
 ||||: |:||  
 Db 187 FQGVVGNVKFV 197

# RESULT 17

## Q8BVV4

ID Q8BVV4 PRELIMINARY; PRT; 750 AA.  
 AC Q8BVV4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Inhibin binding protein long isoform homolog.  
 GN AI747649.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Head;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium,  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK076427; BAC36338.1; -.  
 DR PIR; PT0635; PT0696.  
 DR MGD; MGI:2147913; AI747649.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003598; Ig\_c2.  
 DR Pfam; PF00047; ig; 5.  
 DR SMART; SM00408; IGc2; 4.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 SQ SEQUENCE 750 AA; 82507 MW; A565CC78E5879DAE CRC64;

Query Match 62.7%; Score 37; DB 11; Length 750;  
 Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRF 10  
 |||: | :||  
 Db 309 FQGIFQGMRF 318

# RESULT 18

Q8QG19

ID Q8QG19 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG19;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Menura novaehollandiae (superb lyrebird).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Menuridae; Menura.  
 OX NCBI\_TaxID=47692;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057004; AAL18121.1; -.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.



FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 110018 MW; EEB76DAAB2BDD9E6 CRC64;

Query Match 62.7%; Score 37; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 1.6e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:| ||:|  
Db 662 GILRSFRFIF 671

#### RESULT 19

Q8LTK2

ID Q8LTK2 PRELIMINARY; PRT; 1711 AA.  
AC Q8LTK2;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Putative tail lysin.  
GN ORF43.  
OS Lactococcus bacteriophage 4268.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=201847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Trotter M., Ross R.P., Fitzgerald G., Coffey A.;  
RT "Sequence analysis of an obligately lytic P335 bacteriophage 4268."  
RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF489521; AAM83082.1; -.  
SQ SEQUENCE 1711 AA; 183092 MW; C1F33C5DEC82D96C CRC64;

Query Match 62.7%; Score 37; DB 9; Length 1711;  
Best Local Similarity 45.5%; Pred. No. 2.8e+02;  
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 11  
|||:::|:|:  
Db 701 FQGIVDTVKFI 711

#### RESULT 20

Q94MA1

ID Q94MA1 PRELIMINARY; PRT; 1713 AA.  
AC Q94MA1;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Lactococcus phage BK5-T.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
OX NCBI\_TaxID=31754;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21235794; PubMed=11336549;  
RA Desiere F., Mahanivong C., Hillier A.J., Chandry P.S., Davidson B.E.,

RA Brussow H.;  
 RT "Comparative Genomics of Lactococcal Phages: Insight from the Complete  
 RT Genome Sequence of Lactococcus lactis Phage BK5-T.";  
 RL Virology 283:240-252(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Mahanivong C., Boyce J.D., Davidson B.E., Hillier A.J.;  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Mahanivong C., Boyce J.D., Davidson B.E., Hillier A.J.;  
 RT "Analysis of the sequence, cos site and structural proteins of the  
 RT Lactococcus lactis temperate bacteriophage BK5-T.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF176025; AAK56815.1; -.  
 DR EMBL; AJ245616; CAC80156.1; -.  
 DR InterPro; IPR008258; SLT\_dom.  
 DR Pfam; PF01464; SLT; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 1713 AA; 183493 MW; DADDEB64DF875278 CRC64;

Query Match 62.7%; Score 37; DB 9; Length 1713;  
 Best Local Similarity 45.5%; Pred. No. 2.8e+02;  
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLSVRFV 11  
 |||::|::|:  
 Db 701 FQIVDTVKFI 711

# RESULT 21

Q9ZEN6

ID Q9ZEN6 PRELIMINARY; PRT; 210 AA.  
 AC Q9ZEN6;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 OS Wolinella succinogenes.  
 OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;  
 OC Helicobacteraceae; Wolinella.  
 OX NCBI\_TaxID=844;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20461222; PubMed=11004174;  
 RA Ullmann R., Gross R., Simon J., Udden G., Kroeger A.;  
 RT "Transport of C4-Dicarboxylates in Wolinella succinogenes.";  
 RL J. Bacteriol. 182:5757-5764(2000).  
 DR EMBL; AJ131242; CAA10333.1; -.  
 DR InterPro; IPR000541; UPF0021.  
 DR Pfam; PF01171; ATP\_bind3; 1.  
 KW Hypothetical protein.  
 FT NON\_TER 210 210  
 SQ SEQUENCE 210 AA; 24738 MW; CF04A38C30C0550F CRC64;

Query Match 61.0%; Score 36; DB 2; Length 210;  
 Best Local Similarity 70.0%; Pred. No. 53;

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFV 11  
||:|||| |:|  
Db 201 QGLLQSFRFL 210

RESULT 22

Q8ENF8

ID Q8ENF8 PRELIMINARY; PRT; 219 AA.  
AC Q8ENF8;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Hypothetical conserved protein.  
GN OB2525.  
OS Oceanobacillus iheyensis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Oceanobacillus.  
OX NCBI\_TaxID=182710;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HTE831 / DSM 14371 / JCM 11309;  
RX MEDLINE=22220767; PubMed=12235376;  
RA Takami H., Takaki Y., Uchiyama I.;  
RT "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya  
RT Ridge and its unexpected adaptive capabilities to extreme  
RT environments.";  
RL Nucleic Acids Res. 30:3927-3935(2002).  
DR EMBL; AP004601; BAC14481.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 219 AA; 24364 MW; E38483308E3866FE CRC64;

Query Match 61.0%; Score 36; DB 16; Length 219;  
Best Local Similarity 58.3%; Pred. No. 55;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
|||| | |:|  
Db 139 FQGVSSFSFIF 150

RESULT 23

Q90XK6

ID Q90XK6 PRELIMINARY; PRT; 325 AA.  
AC Q90XK6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE RAG1 (Fragment).  
GN RAG1.  
OS Amia calva (Bowfin).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Amiiformes; Amiidae; Amia.  
OX NCBI\_TaxID=7924;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21457245; PubMed=11553795;

RA Venkatesh B., Erdmann M.V., Brenner S.;  
 RT "Molecular synapomorphies resolve evolutionary relationships of extant  
 RT jawed vertebrates."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:11382-11387(2001).  
 DR EMBL; AF369059; AAL12861.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 325 325  
 SQ SEQUENCE 325 AA; 36432 MW; 3CF8B6B7667BE722 CRC64;

Query Match 61.0%; Score 36; DB 13; Length 325;  
 Best Local Similarity 70.0%; Pred. No. 82;  
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:| ||||  
 Db 217 GMLRSFRFVF 226

#### RESULT 24

Q87M68

ID Q87M68 PRELIMINARY; PRT; 394 AA.  
 AC Q87M68;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Conserved hypothetical protein.  
 GN VP2390.  
 OS Vibrio parahaemolyticus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=670;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;  
 RX MEDLINE=22508454; PubMed=12620739;  
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
 RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism  
 RT distinct from that of V. cholerae."  
 RL Lancet 361:743-749(2003).  
 DR EMBL; AP005081; BAC60653.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 394 AA; 44540 MW; 341AC16555A97B02 CRC64;

Query Match 61.0%; Score 36; DB 16; Length 394;  
 Best Local Similarity 72.7%; Pred. No. 1e+02;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 11  
 | || | ||||  
 Db 246 FAGVAQLVRFV 256

#### RESULT 25

Q8U2Q0

ID Q8U2Q0 PRELIMINARY; PRT; 405 AA.

AC Q8U2Q0;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE Hypothetical protein PF0783.  
 GN PF0783.  
 OS Pyrococcus furiosus.  
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;  
 OC Pyrococcus.  
 OX NCBI\_TaxID=2261;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;  
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;  
 RT "The complete sequence of the Pyrococcus furiosus genome.";  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AE010195; AAL80907.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 405 AA; 45094 MW; 0D206BC87EA8F312 CRC64;  
  
 Query Match 61.0%; Score 36; DB 17; Length 405;  
 Best Local Similarity 63.6%; Pred. No. 1e+02;  
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
 ||:| ::|||  
 Db 156 QGLLFALRFVF 166

# RESULT 26

Q9UEU4  
 ID Q9UEU4 PRELIMINARY; PRT; 595 AA.  
 AC Q9UEU4;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Imidazoline receptor antisera-selected protein (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99066883; PubMed=9851558;  
 RA Ivanov T.R., Jones J.C., Dontenwill M., Bousquet P., Piletz J.E.;  
 RT "Characterization of a partial cDNA clone detected by imidazoline  
 RT receptor-selective antisera.";  
 RL J. Auton. Nerv. Syst. 72:98-110(1998).  
 DR EMBL; AF058290; AAC33321.1; -.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 KW Receptor.  
 FT NON\_TER 1 1  
 FT NON\_TER 595 595  
 SQ SEQUENCE 595 AA; 65354 MW; 93139B536F447CBD CRC64;

Query Match 61.0%; Score 36; DB 4; Length 595;  
Best Local Similarity 63.6%; Pred. No. 1.5e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
| :| |:||||  
Db 283 QHILSSLRFVF 293

RESULT 27

Q8QG47

ID Q8QG47 PRELIMINARY; PRT; 956 AA.  
AC Q8QG47;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Aegithalos iouschensis.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Aegithalidae;  
OC Aegithalos.  
OX NCBI\_TaxID=174998;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY056976; AAL18093.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 956 956  
SQ SEQUENCE 956 AA; 109831 MW; CA13CF9F4E35116A CRC64;

Query Match 61.0%; Score 36; DB 13; Length 956;  
Best Local Similarity 70.0%; Pred. No. 2.5e+02;  
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|||:: ||||  
Db 661 GVLRTFRFVF 670

RESULT 28

Q9UT41

ID Q9UT41 PRELIMINARY; PRT; 1142 AA.  
 AC Q9UT41;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Putative negative regulator of vesicle formation.  
 GN SPAC824.02.  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972h-;  
 RA Barrell B.G., Rajandream M.A., Quail M., Seegar K., Harris D.;  
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL121741; CAB57332.1; -.  
 DR PIR; T39103; T39103.  
 DR GeneDB\_SPombe; SPAC824.02; -.  
 DR GO; GO:0003824; F:catalytic activity; IEA.  
 DR InterPro; IPR008262; Lipase\_AS.  
 DR InterPro; IPR000379; Ser\_estrs.  
 DR PROSITE; PS00120; LIPASE\_SER; 1.  
 SQ SEQUENCE 1142 AA; 129412 MW; A550AB2916DA8C89 CRC64;

Query Match 61.0%; Score 36; DB 3; Length 1142;  
 Best Local Similarity 58.3%; Pred. No. 3e+02;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 FQGVLSQSVRFVF 12  
 | || |:|:| |  
 Db 887 FTGVCQAVKFSF 898

RESULT 29

Q9EPW8

ID Q9EPW8 PRELIMINARY; PRT; 1354 AA.  
 AC Q9EPW8;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Nischarin.  
 GN NISCH.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ICR outbred; TISSUE=Brain;  
 RX MEDLINE=20571837; PubMed=11121431;  
 RA Alahari S.K., Lee J.W., Juliano R.L.;  
 RT "Nischarin, a Novel Protein That Interacts with the Integrin alpha5  
 RT Subunit and Inhibits Cell Migration.";

RL J. Cell Biol. 151:1141-1154(2000).  
 DR EMBL; AF315344; AAG42100.1; -.  
 DR MGD; MGI:1928323; Nisch.  
 DR GO; GO:0005829; C:cytosol; IDA.  
 DR GO; GO:0005178; F:integrin binding; IDA.  
 DR GO; GO:0030036; P:actin cytoskeleton organization and biogenesis; IDA.  
 DR GO; GO:0030336; P:negative regulation of cell migration; IDA.  
 DR GO; GO:0016601; P:RAC protein signal transduction; IDA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR007092; LRR\_SDS22.  
 DR Pfam; PF00560; LRR; 4.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 SQ SEQUENCE 1354 AA; 148060 MW; 01BD676FDCA19247 CRC64;

Query Match 61.0%; Score 36; DB 11; Length 1354;  
 Best Local Similarity 63.6%; Pred. No. 3.5e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLSVRFVF 12  
 | : | | : | | |  
 Db 506 QHILSSLRFVF 516

# RESULT 30

## Q80TM9

ID Q80TM9 PRELIMINARY; PRT; 1480 AA.  
 AC Q80TM9;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE MKIAA0975 protein (Fragment).  
 GN MKIAA0975.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=22579291; PubMed=12693553;  
 RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Aizawa H., Yuasa S.,  
 RA Nakajima D., Nagase T., Ohara O., Koga H.;  
 RT "Prediction of the coding sequences of mouse homologues of KIAA gene:  
 RT II. The complete nucleotide sequences of 400 mouse KIAA-homologous  
 RT cDNAs identified by screening of terminal sequences of cDNA clones  
 RT randomly sampled from size-fractionated libraries."  
 RL DNA Res. 10:35-48(2003).  
 DR EMBL; AK122412; BAC65694.1; -.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003885; LRR\_cyst.  
 DR InterPro; IPR001683; PX.  
 DR Pfam; PF00560; LRR; 3.



DR Pfam; PF00787; PX; 1.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00365; LRR\_SD22; 3.  
 DR SMART; SM00312; PX; 1.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PROSITE; PS50195; PX; 1.  
 FT NON\_TER 1 1  
 SQ SEQUENCE 1480 AA; 162411 MW; 0B4A907A5156BB6C CRC64;

Query Match 61.0%; Score 36; DB 11; Length 1480;  
 Best Local Similarity 63.6%; Pred. No. 3.9e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
 | : | | : | | |  
 Db 644 QHILSSLRFVF 654

# RESULT 31

## Q9UES6

ID Q9UES6 PRELIMINARY; PRT; 1504 AA.  
 AC Q9UES6; Q9UFW3;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE I-1 receptor candidate protein.  
 GN DKFZP434E066.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20338079; PubMed=10882231;  
 RA Piletz J.E., Ivanov T.R., Sharp J.D., Ernsberger P., Chang C.H.,  
 RA Pickard R.T., Gold G., Roth B., Zhu H., Jones J.C., Baldwin J.,  
 RA Reis D.J.;  
 RT "Imidazoline receptor antisera-selected (IRAS) cDNA: cloning and  
 RT characterization.";  
 RL DNA Cell Biol. 19:319-329(2000).  
 RN [2]  
 RP SEQUENCE OF 512-1504 FROM N.A.  
 RC TISSUE=Testis;  
 RA Koehrer K., Beyer A., Mewes H.W., Gassenhuber J., Wiemann S.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF082516; AAC33104.1; -.  
 DR EMBL; AL117432; CAB55920.1; -.  
 DR PIR; T17230; T17230.  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR InterPro; IPR001128; Cytochrome\_P450.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR007092; LRR\_SDS22.  
 DR InterPro; IPR001683; PX.  
 DR Pfam; PF00560; LRR; 5.  
 DR Pfam; PF00787; PX; 1.

DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00312; PX; 1.  
 DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
 DR PROSITE; PS50195; PX; 1.  
 KW Hypothetical protein; Receptor.  
 SQ SEQUENCE 1504 AA; 166670 MW; 4430726CB4DE959F CRC64;

Query Match 61.0%; Score 36; DB 4; Length 1504;  
 Best Local Similarity 63.6%; Pred. No. 3.9e+02;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
 | : | | : | | |  
 Db 751 QHILSSLRFVF 761

# RESULT 32

Q7Z2X6

ID Q7Z2X6 PRELIMINARY; PRT; 1504 AA.  
 AC Q7Z2X6;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE NISCH protein.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Peripheral Nervous System;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Peripheral Nervous System;  
 RA Strausberg R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DDBJ databases.

DR EMBL; BC054494; AAH54494.1; -.  
SQ SEQUENCE 1504 AA; 166675 MW; F62ABEDF4B19C965 CRC64;

Query Match 61.0%; Score 36; DB 4; Length 1504;  
Best Local Similarity 63.6%; Pred. No. 3.9e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QGVLQSVRFVF 12  
| :| | :|||  
Db 751 QHILSSLRFVF 761

# RESULT 33

## Q9Y2I1

ID Q9Y2I1 PRELIMINARY; PRT; 1528 AA.  
AC Q9Y2I1;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein KIAA0975 (Fragment).  
GN KIAA0975.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99246063; PubMed=10231032;  
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,  
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;  
RT "Prediction of the coding sequences of unidentified human genes. XIII.  
RT The complete sequences of 100 new cDNA clones from brain which code  
RT for large proteins in vitro.";  
RL DNA Res. 6:63-70(1999).  
DR EMBL; AB023192; BAA76819.1; -.  
DR Genew; HGNC:18006; NISCH.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
DR InterPro; IPR001128; Cytochrome\_P450.  
DR InterPro; IPR001611; LRR.  
DR InterPro; IPR007092; LRR\_SDS22.  
DR InterPro; IPR001683; PX.  
DR Pfam; PF00560; LRR; 5.  
DR Pfam; PF00787; PX; 1.  
DR PRINTS; PR00019; LEURICHRPT.  
DR SMART; SM00312; PX; 1.  
DR PROSITE; PS00086; CYTOCHROME\_P450; 1.  
DR PROSITE; PS50195; PX; 1.  
KW Hypothetical protein.  
FT NON\_TER 1 1  
SQ SEQUENCE 1528 AA; 168840 MW; FEB1DADD65787A5B CRC64;

Query Match 61.0%; Score 36; DB 4; Length 1528;  
Best Local Similarity 63.6%; Pred. No. 4e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
| : | | : | | |  
Db 775 QHILSSLRFVF 785

RESULT 34

Q7YY16

ID Q7YY16 PRELIMINARY; PRT; 1554 AA.  
AC Q7YY16;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN 1MX.02.  
OS Cryptosporidium parvum.  
OC Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida;  
OC Cryptosporidiidae; Cryptosporidium.  
OX NCBI\_TaxID=5807;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Iowa;  
RA Bankier A.T., Spriggs H.F., Fartmann B., Konfortov B.A., Madera M.,  
RA Vogel C., Teichmann S.A., Ivens A., Dear P.H.;  
RT "Integrated mapping, chromosomal sequencing and sequence analysis of  
RT Cryptosporidium parvum.";  
RL Genome Res. 0:0-0(2003).  
DR EMBL; BX538350; CAD98672.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 1554 AA; 181712 MW; 03853ED293A8626C CRC64;

Query Match 60.2%; Score 35.5; DB 5; Length 1554;  
Best Local Similarity 66.7%; Pred. No. 5.1e+02;  
Matches 8; Conservative 3; Mismatches 0; Indels 1; Gaps 1;

Qy 1 FQGVLSVRFVF 12  
| | : | | | : : | | |  
Db 1328 FQ-LLQSIKFVF 1338

RESULT 35

Q9ZUI2

ID Q9ZUI2 PRELIMINARY; PRT; 82 AA.  
AC Q9ZUI2;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE T2K10.16 protein.  
GN T2K10.16.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Vysotskaia V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,  
RA Li J., Kremenetskaia I., Luros J., Gonzalez A., Altafi H., Araujo R.,

RA Buehler E., Chao Q., Conn L., Conway A.B., Dunn P., Hansen N.,  
 RA Huizar L., Kim C., Palm C., Rowley D., Shinn P., Walker M.,  
 RA Davis R.W., Ecker J.R., Federspiel N.A., Theologis A.;  
 RT "Arabidopsis thaliana chromosome 1 BAC T2K10 sequence."  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC005966; AAD14486.1; -.  
 DR PIR; F96625; F96625.  
 DR InterPro; IPR001229; Jacalin\_lectin.  
 DR Pfam; PF01419; Jacalin; 1.  
 SQ SEQUENCE 82 AA; 9478 MW; 0461CE1F3C584141 CRC64;

Query Match 59.3%; Score 35; DB 10; Length 82;  
 Best Local Similarity 63.6%; Pred. No. 32;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 11  
 |||:|:| |  
 Db 69 FQGLLQTVSLV 79

# RESULT 36

## Q8RC20

ID Q8RC20 PRELIMINARY; PRT; 217 AA.  
 AC Q8RC20;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE F0F1-type ATP synthase a subunit.  
 GN ATPB OR TTE0630.  
 OS Thermoanaerobacter tengcongensis.  
 OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;  
 OC Thermoanaerobacteriaceae; Thermoanaerobacter.  
 OX NCBI\_TaxID=119072;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MB4 / JCM 11007;  
 RX MEDLINE=21992816; PubMed=11997336;  
 RA Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,  
 RA Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,  
 RA Tan H., Chen R., Wang J., Yu J., Yang H.;  
 RT "A complete sequence of T. tengcongensis genome."  
 RL Genome Res. 12:689-700(2002).  
 DR EMBL; AE013032; AAM23899.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0015992; P:proton transport; IEA.  
 DR InterPro; IPR000568; ATPsynt\_Asub.  
 DR Pfam; PF00119; ATP-synt\_A; 1.  
 DR PRINTS; PR00123; ATPASEA.  
 DR TIGRFAMS; TIGR01131; ATP\_synt\_6\_or\_A; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 217 AA; 23895 MW; 5AD4467861AE7EEE CRC64;

Query Match 59.3%; Score 35; DB 16; Length 217;  
 Best Local Similarity 58.3%; Pred. No. 86;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 12

Db                   | | : | | : | |  
195 FDGLLQMLVFVF 206

RESULT 37

Q7ZWN5

ID   Q7ZWN5           PRELIMINARY;           PRT;   248 AA.  
AC   Q7ZWN5;  
DT   01-JUN-2003 (TrEMBLrel. 24, Created)  
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE   Similar to triosephosphate isomerase.  
OS   Xenopus laevis (African clawed frog).  
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC   Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
OC   Xenopodinae; Xenopus.  
OX   NCBI\_TaxID=8355;  
RN   [1]  
RP   SEQUENCE FROM N.A.  
RC   TISSUE=Embryo;  
RA   Klein S., Strausberg R.;  
RL   Submitted (FEB-2003) to the EMBL/GenBank/DDBJ databases.  
DR   EMBL; BC046864; AAH46864.1; -.  
DR   GO; GO:0016853; F:isomerase activity; IEA.  
DR   GO; GO:0004807; F:triose-phosphate isomerase activity; IEA.  
DR   GO; GO:0008152; P:metabolism; IEA.  
DR   InterPro; IPR000652; Triophos\_ismrse.  
DR   Pfam; PF00121; TIM; 1.  
DR   ProDom; PD001005; Triophos\_ismrse; 1.  
DR   TIGRFAMs; TIGR00419; tim; 1.  
DR   PROSITE; PS00171; TIM; 1.  
KW   Isomerase.  
SQ   SEQUENCE   248 AA;   26762 MW;   2EC929CD0CF6E804 CRC64;  
  
Query Match                   59.3%;   Score 35;   DB 13;   Length 248;  
Best Local Similarity       54.5%;   Pred. No. 99;  
Matches       6;   Conservative    3;   Mismatches    2;   Indels     0;   Gaps     0;

Qy           2 QGVLQSVRFVF 12  
             : | | | | : :  
Db           198 EGVAQSVRIIY 208

RESULT 38

Q90XK7

ID   Q90XK7           PRELIMINARY;           PRT;   323 AA.  
AC   Q90XK7;  
DT   01-DEC-2001 (TrEMBLrel. 19, Created)  
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE   RAG1 (Fragment).  
GN   RAG1.  
OS   Lepisosteus osseus (Long-nosed gar).  
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC   Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;  
OC   Lepisosteus.  
OX   NCBI\_TaxID=34771;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21457245; PubMed=11553795;  
 RA Venkatesh B., Erdmann M.V., Brenner S.;  
 RT "Molecular synapomorphies resolve evolutionary relationships of extant  
 RT jawed vertebrates."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:11382-11387(2001).  
 DR EMBL; AF369058; AAL12860.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 323 323  
 SQ SEQUENCE 323 AA; 36189 MW; 0CFA2A21B34E434D CRC64;

Query Match 59.3%; Score 35; DB 13; Length 323;  
 Best Local Similarity 60.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |::| |::|  
 Db 216 GLLRSFRFIF 225

# RESULT 39

## Q94DF6

ID Q94DF6 PRELIMINARY; PRT; 460 AA.  
 AC Q94DF6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE P0518C01.32 protein (P0698H10.10 protein).  
 GN P0518C01.32 OR P0698H10.10.  
 OS Oryza sativa (Rice), and  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530, 39947;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0518C01."  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Nipponbare;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:P0698H10."  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AP003277; BAB63693.1; -.  
 DR EMBL; AP003298; BAB89867.1; -.  
 DR Gramene; Q94DF6; -.  
 SQ SEQUENCE 460 AA; 50542 MW; F716E85DD61B00DF CRC64;

Query Match 59.3%; Score 35; DB 10; Length 460;  
 Best Local Similarity 77.8%; Pred. No. 1.9e+02;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 VLQSVRFVF 12  
||| |:|||  
Db 173 VLQGVKFVF 181

RESULT 40

Q8ZCV8

ID Q8ZCV8 PRELIMINARY; PRT; 465 AA.  
AC Q8ZCV8;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative membrane protein (Putative permease, major facilitator  
DE superfamily).  
GN YEGB OR YPO2850 OR Y1383.  
OS Yersinia pestis.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Yersinia.  
OX NCBI\_TaxID=632;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CO-92 / Biovar Orientalis;  
RX MEDLINE=21470413; PubMed=11586360;  
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,  
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,  
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,  
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,  
RA Feltwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,  
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,  
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;  
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";  
RL Nature 413:523-527(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=KIM5 / Biovar Mediaevalis;  
RX MEDLINE=22137863; PubMed=12142430;  
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
RA Perry R.D.;  
RT "Genome sequence of Yersinia pestis KIM.";  
RL J. Bacteriol. 184:4601-4611(2002).  
DR EMBL; AJ414154; CAC92102.1; -.  
DR EMBL; AE013741; AAM84955.1; -.  
DR PIR; AC0347; AC0347.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0015904; P:tetracycline transport; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR001411; TCR\_TetB.  
DR Pfam; PF00083; sugar\_tr; 1.



DR PRINTS; PR01036; TCRTETB.  
DR PROSITE; PS50850; MFS; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 465 AA; 50176 MW; 0CC273F10BB3F5ED CRC64;

Query Match 59.3%; Score 35; DB 16; Length 465;  
Best Local Similarity 60.0%; Pred. No. 1.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLSVRF 10  
|||:| ::||  
Db 362 FQGMLNALRF 371

#### RESULT 41

Q8GH66

ID Q8GH66 PRELIMINARY; PRT; 593 AA.  
AC Q8GH66;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative ABC transporter.  
OS Mycobacterium paratuberculosis.  
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1770;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strommenger B., Gerlach G.F.;  
RT "Putative ABC transporter."  
RL Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AF419325; AAN77618.1; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.  
DR GO; GO:0000166; F:nucleotide binding; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR001140; ABC\_TM\_transpt.  
DR InterPro; IPR003439; ABC\_transporter.  
DR Pfam; PF00664; ABC\_membrane; 1.  
DR Pfam; PF00005; ABC\_tran; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC\_TRANSPORTER\_1; 1.  
DR PROSITE; PS50893; ABC\_TRANSPORTER\_2; 1.  
SQ SEQUENCE 593 AA; 63847 MW; E155A821B8A63C19 CRC64;

Query Match 59.3%; Score 35; DB 2; Length 593;  
Best Local Similarity 54.5%; Pred. No. 2.4e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLSVRFVF 12  
||:|| : |:|  
Db 30 QGILQGLGFLF 40

#### RESULT 42

Q822D4

ID Q822D4 PRELIMINARY; PRT; 610 AA.  
AC Q822D4;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ABC transporter, permease protein, putative.  
GN CCA00749.  
OS Chlamydophila caviae.  
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.  
OX NCBI\_TaxID=83557;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GPIC;  
RX MEDLINE=22569155; PubMed=12682364;  
RA Read T.D., Myers G.S.A., Brunham R.C., Nelson W.C., Paulsen I.T.,  
RA Heidelberg J., Holtzapple E., Khouri H., Federova N.B., Carty H.A.,  
RA Umayam L.A., Haft D.H., Peterson J., Beanan M.J., White O.,  
RA Salzberg S.L., Hsia R.-C., McClarty G., Rank R.G., Bavoil P.M.,  
RA Fraser C.M.;  
RT "Genome sequence of Chlamydophila caviae (Chlamydia psittaci GPIC):  
RT examining the role of niche-specific genes in the evolution of the  
RT Chlamydiaceae.";  
RL Nucleic Acids Res. 31:2134-2147(2003).  
DR EMBL; AE016996; AAP05490.1; -.  
DR TIGR; CCA00749; -.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR000515; BPD\_transp.  
DR Pfam; PF00528; BPD\_transp; 1.  
KW Complete proteome.  
SQ SEQUENCE 610 AA; 68001 MW; B64FD0153AD1A49D CRC64;

Query Match 59.3%; Score 35; DB 16; Length 610;  
Best Local Similarity 50.0%; Pred. No. 2.5e+02;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
|||:| |:|  
Db 231 FQGILLSEKLIF 242

RESULT 43

Q87GG1

ID Q87GG1 PRELIMINARY; PRT; 627 AA.  
AC Q87GG1;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Putative type III secretion system EscV protein.  
GN VPA1355.  
OS Vibrio parahaemolyticus.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
OC Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=670;  
RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=RIMD 2210633 / Serotype O3:K6;  
 RX MEDLINE=22508454; PubMed=12620739;  
 RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,  
 RA Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,  
 RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;  
 RT "Genome sequence of *Vibrio parahaemolyticus*: a pathogenic mechanism  
 RT distinct from that of *V. cholerae*.";  
 RL Lancet 361:743-749(2003).  
 DR EMBL; AP005088; BAC62698.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0009306; P:protein secretion; IEA.  
 DR InterPro; IPR001712; Bact\_FHIPEP.  
 DR Pfam; PF00771; FHIPEP; 1.  
 DR PRINTS; PR00949; TYPE3IMAPROT.  
 KW Complete proteome.  
 SQ SEQUENCE 627 AA; 69902 MW; 3A3044A6695AE092 CRC64;

Query Match 59.3%; Score 35; DB 16; Length 627;  
 Best Local Similarity 60.0%; Pred. No. 2.6e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 | : | | : | | |  
 Db 337 GLFQSLKFVF 346

#### RESULT 44

##### Q8MNS5

ID Q8MNS5 PRELIMINARY; PRT; 657 AA.  
 AC Q8MNS5;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN F42G9.6.  
 OS *Caenorhabditis elegans*.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; *Caenorhabditis*.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Waterston R.;  
 RT "Genome sequence of the nematode *C. elegans*: a platform for  
 RT investigating biology. The *C. elegans* Sequencing Consortium.";  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA R Wilson., Taich A.;  
 RT "The sequence of *C. elegans* cosmid F42G9.";  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;

RA Waterston R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; U00051; AAM29694.1; -.  
 DR WormPep; F42G9.6a; CE30791.  
 DR GO; GO:0004806; F:triacylglycerol lipase activity; IEA.  
 DR GO; GO:0006629; P:lipid metabolism; IEA.  
 DR InterPro; IPR002921; Lipase\_3.  
 DR Pfam; PF01764; Lipase\_3; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 657 AA; 73751 MW; 56F8595EEBE514AC CRC64;

Query Match 59.3%; Score 35; DB 5; Length 657;  
 Best Local Similarity 54.5%; Pred. No. 2.7e+02;  
 Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
 :|:|:| |:|  
 Db 395 RGMLRSARYVF 405

#### RESULT 45

##### Q8MNS4

ID Q8MNS4 PRELIMINARY; PRT; 683 AA.  
 AC Q8MNS4;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN F42G9.6.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA Waterston R.;  
 RT "Genome sequence of the nematode C. elegans: a platform for  
 RT investigating biology. The C. elegans Sequencing Consortium."  
 RL Science 282:2012-2018(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA R Wilson., Taich A.;  
 RT "The sequence of C. elegans cosmid F42G9."  
 RL Submitted (MAY-1994) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Waterston R.;  
 RL Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; U00051; AAM29695.1; -.  
 DR WormPep; F42G9.6b; CE30792.  
 DR GO; GO:0004806; F:triacylglycerol lipase activity; IEA.  
 DR GO; GO:0006629; P:lipid metabolism; IEA.  
 DR InterPro; IPR002921; Lipase\_3.

DR Pfam; PF01764; Lipase\_3; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 683 AA; 76773 MW; 980BAF8D757C8BCB CRC64;

Query Match 59.3%; Score 35; DB 5; Length 683;  
Best Local Similarity 54.5%; Pred. No. 2.8e+02;  
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 2 QGVLSQSVRFVF 12  
:|:|:| |:|  
Db 421 RGMRLRSARYVF 431

#### RESULT 46

Q9V780

ID Q9V780 PRELIMINARY; PRT; 849 AA.  
AC Q9V780;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE CG10255 protein.  
GN LAP1 OR CG10255.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 CC -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.  
 DR EMBL; AE003812; AAF58179.1; -.  
 DR FlyBase; FBgn0033984; Lap1.  
 DR GO; GO:0007242; P:intracellular signaling cascade; IEA.  
 DR InterPro; IPR001611; LRR.  
 DR InterPro; IPR003591; LRR\_typ.  
 DR InterPro; IPR001478; PDZ.  
 DR Pfam; PF00560; LRR; 10.  
 DR Pfam; PF00595; PDZ; 1.  
 DR PRINTS; PR00019; LEURICHRPT.  
 DR SMART; SM00369; LRR\_TYP; 1.  
 DR SMART; SM00228; PDZ; 1.  
 DR PROSITE; PS50106; PDZ; 1.  
 SQ SEQUENCE 849 AA; 96147 MW; C86DD96C242080E2 CRC64;

Query Match 59.3%; Score 35; DB 5; Length 849;  
 Best Local Similarity 60.0%; Pred. No. 3.5e+02;  
 Matches 6; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|:|:|:|:  
 Db 313 GMLRSLRFLF 322

# RESULT 47

## Q8QFY3

ID Q8QFY3 PRELIMINARY; PRT; 951 AA.  
 AC Q8QFY3;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Vanga curvirostris.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;  
 OC Malaconotinae; Vanga.  
 OX NCBI\_TaxID=125301;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";

RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057040; AAL18157.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 951 951  
 SQ SEQUENCE 951 AA; 109162 MW; 8B385054E4E1C304 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 951;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 656 GILRTFRFVF 665

# RESULT 48

## Q8QFY5

ID Q8QFY5 PRELIMINARY; PRT; 952 AA.  
 AC Q8QFY5;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Troglodytes aedon (House wren).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Certhiidae;  
 OC Troglodytinae; Troglodytes.  
 OX NCBI\_TaxID=58211;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057038; AAL18155.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.

DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 952 952  
 SQ SEQUENCE 952 AA; 108913 MW; E38DE40C8FAE3A61 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 952;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 657 GILRTFRFVF 666

# RESULT 49

Q8QFY9

ID Q8QFY9 PRELIMINARY; PRT; 952 AA.  
 AC Q8QFY9;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Thamnophilus nigrocinereus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Thamnophilidae;  
 OC Thamnophilus.  
 OX NCBI\_TaxID=175015;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057034; AAL18151.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 952 952  
 SQ SEQUENCE 952 AA; 109127 MW; 5E53915E22D4BD76 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 952;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;



Qy 3 GVLQSVRFVF 12  
|:|::|  
Db 657 GILRTFRFVF 666

RESULT 50

Q8QG14

ID Q8QG14 PRELIMINARY; PRT; 953 AA.  
AC Q8QG14;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Nectarinia olivacea.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
OC Nectariniidae; Nectarinia.  
OX NCBI\_TaxID=170863;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057009; AAL18126.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 953 953  
SQ SEQUENCE 953 AA; 109284 MW; 003D5BC35D390475 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 953;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|::|  
Db 658 GILRTFRFVF 667

RESULT 51

Q8QG08

ID Q8QG08 PRELIMINARY; PRT; 955 AA.  
AC Q8QG08;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Pardalotus striatus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Pardalotidae;  
 OC Pardalotus.  
 OX NCBI\_TaxID=108840;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057015; AAL18132.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 955 955  
 SQ SEQUENCE 955 AA; 109602 MW; E9511BB60678C70E CRC64;

Query Match 59.3%; Score 35; DB 13; Length 955;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|:: |||  
 Db 660 GILRTFRFVF 669

# RESULT 52

Q8QFY7

ID Q8QFY7 PRELIMINARY; PRT; 956 AA.  
 AC Q8QFY7;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Toxorhamphus novaeguineae.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
 OC Melanocharitidae; Toxorhamphus.  
 OX NCBI\_TaxID=175137;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057036; AAL18153.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 956 956  
 SQ SEQUENCE 956 AA; 109368 MW; D430D49CE2774036 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 956;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |::: |||  
 Db 661 GILRTFRFVF 670

# RESULT 53

Q9W6P9

ID Q9W6P9 PRELIMINARY; PRT; 956 AA.  
 AC Q9W6P9;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Tyrannus tyrannus (Eastern kingbird).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Tyrannidae; Tyrannus.  
 OX NCBI\_TaxID=43165;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99310776; PubMed=10381315;  
 RA Groth J.G., Barrowclough G.F.;  
 RT "Basal divergences in birds and the phylogenetic utility of the  
 RT nuclear RAG-1 gene.";  
 RL Mol. Phylogenet. Evol. 12:115-123(1999).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AF143739; AAD34965.1; -.  
 DR HSSP; P15919; 1RMD.  
 DR InterPro; IPR001841; Znf\_ring.

DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 956 956  
 SQ SEQUENCE 956 AA; 109688 MW; C3DBC86167D4B280 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 956;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 661 GILRTFRFVF 670

# RESULT 54

## Q8QFZ5

ID Q8QFZ5 PRELIMINARY; PRT; 957 AA.  
 AC Q8QFZ5;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS *Regulus calendula* (Ruby-crowned kinglet).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Sylviidae; *Regulus*.  
 OX NCBI\_TaxID=73321;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057028; AAL18145.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109716 MW; 6BC1179B98084D3C CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;

Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|::|  
Db 662 GILRTFRFVF 671

RESULT 55

Q8QG23

ID Q8QG23 PRELIMINARY; PRT; 957 AA.  
AC Q8QG23;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Lanius ludovicianus (Loggerhead shrike).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Laniidae;  
OC Lanius.  
OX NCBI\_TaxID=28713;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057000; AAL18117.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS0089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109729 MW; A037E9DCB05DABB7 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|::|  
Db 662 GILRTFRFVF 671

RESULT 56

Q8QFZ4

ID Q8QFZ4 PRELIMINARY; PRT; 957 AA.

AC Q8QFZ4;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Rupicola rupicola.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Cotingidae; Rupicola.  
 OX NCBI\_TaxID=81879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057029; AAL18146.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109909 MW; 12E0C5D50C314709 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |::: |||  
 Db 662 GILRTFRFVF 671

# RESULT 57

Q8QG38

ID Q8QG38 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG38;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Cinclus cinclus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Cinclidae; Cinclus.  
 OX NCBI\_TaxID=127875;  
 RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY056985; AAL18102.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109741 MW; AF78FE024A9E60C2 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: |||  
Db 662 GILRTFRFVF 671

# RESULT 58

Q8QG12  
ID Q8QG12 PRELIMINARY; PRT; 957 AA.  
AC Q8QG12;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Oriolus larvatus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;  
OC Corvinae; Oriolus.  
OX NCBI\_TaxID=175010;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057011; AAL18128.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.

DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109660 MW; 670C26123BB1AC10 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 59

Q8QG37

ID Q8QG37 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG37;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Cisticola anonymus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Sylviidae; Cisticola.  
 OX NCBI\_TaxID=174999;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056986; AAL18103.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957



SQ SEQUENCE 957 AA; 109872 MW; 9F7BCCACE299F3A4 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 662 GILRTFRFVF 671

RESULT 60

Q8QG10

ID Q8QG10 PRELIMINARY; PRT; 957 AA.  
AC Q8QG10;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Pachycephala soror (Sclater's whistler).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Pachycephalidae;  
OC Pachycephala.  
OX NCBI\_TaxID=156164;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057013; AAL18130.1; -.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109482 MW; 6935F09EACB2E113 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 662 GILRTFRFVF 671

RESULT 61

Q8QFZ1

ID Q8QFZ1 PRELIMINARY; PRT; 957 AA.

AC Q8QFZ1;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Sturnus vulgaris (Starling).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Sturnidae; Sturnus.  
 OX NCBI\_TaxID=9172;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057032; AAL18149.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109523 MW; 0322358C724FC536 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |::: |||  
 Db 662 GILRTFRFVF 671

# RESULT 62

Q8QG04  
 ID Q8QG04 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG04;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Picathartes gymnocephalus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Picathartidae;  
 OC Picathartes.  
 OX NCBI\_TaxID=175131;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057019; AAL18136.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109868 MW; B268695246AF004A CRC64;  
  
 Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
  
 QY 3 GVLQSVRFVF 12  
 |::: ||||  
 Db 662 GILRTFRFVF 671

RESULT 63

Q8QG46

ID Q8QG46 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG46;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Aegithina tiphia.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;  
 OC Aegithininae; Aegithina.  
 OX NCBI\_TaxID=175021;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056977; AAL18094.1; -.

DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109892 MW; 8299C4EAE8450D64 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

RESULT 64

Q8QG15

ID Q8QG15 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG15;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Muscicapa strophiiata.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Muscicapidae; Muscicapa.  
 OX NCBI\_TaxID=175009;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057008; AAL18125.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109662 MW; 54D9B388E32E1A53 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;

Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|::|  
Db 662 GILRTFRFVF 671

RESULT 65

Q8QG26

ID Q8QG26 PRELIMINARY; PRT; 957 AA.  
AC Q8QG26;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Hirundo pyrrhonota.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Hirundinidae; Hirundo.  
OX NCBI\_TaxID=175005;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY056997; AAL18114.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109930 MW; 5C00DE5C38B25D6B CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|::|  
Db 662 GILRTFRFVF 671

RESULT 66

Q8QG28

ID Q8QG28 PRELIMINARY; PRT; 957 AA.

AC Q8QG28;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Furnarius rufus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Furnariidae; Furnarius.  
 OX NCBI\_TaxID=81886;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056995; AAL18112.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS0089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109899 MW; 7C99A6571CFFCDBA CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 67

Q8QG24  
 ID Q8QG24 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG24;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Irena cyanogaster.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Irenidae;  
 OC Irena.  
 OX NCBI\_TaxID=175120;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056999; AAL18116.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109434 MW; 54B905B532660FA9 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy. 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 68

Q8QG21

ID Q8QG21 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG21;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Melanocharis nigra.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
 OC Melanocharitidae; Melanocharis.  
 OX NCBI\_TaxID=175124;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057002; AAL18119.1; -.

DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109641 MW; 9D24D29BCDCF76F5 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|::|||  
 Db 662 GILRTFRFVF 671

# RESULT 69

Q8QG27

ID Q8QG27 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG27;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Garrulax milleti.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Sylviidae; Garrulax.  
 OX NCBI\_TaxID=175004;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056996; AAL18113.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1



FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109379 MW; 5D6BEBF07293F0BE CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: |||  
Db 662 GILRTFRFVF 671

#### RESULT 70

##### Q8QG43

ID Q8QG43 PRELIMINARY; PRT; 957 AA.  
AC Q8QG43;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Artamus leucorhynchus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;  
OC Corvinae; Artamus.  
OX NCBI\_TaxID=175024;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY056980; AAL18097.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109712 MW; EE6F4035949E371F CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: |||  
Db 662 GILRTFRFVF 671

RESULT 71

Q8QFZ0

ID Q8QFZ0 PRELIMINARY; PRT; 957 AA.  
 AC Q8QFZ0;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Sylvia nana.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Sylviidae; Sylvia.  
 OX NCBI\_TaxID=175014;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057033; AAL18150.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109590 MW; E821D215EE1A219D CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: |||  
 Db 662 GILRTFRFVF 671

RESULT 72

Q8QG13

ID Q8QG13 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG13;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.

OS Oedistoma iliolophum.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
 OC Melanocharitidae; Oedistoma.  
 OX NCBI\_TaxID=175128;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057010; AAL18127.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS0089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109572 MW; 40D4CA606917267C CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |::: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 73

Q8QG30

ID Q8QG30 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG30;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Formicarius colma.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Formicariidae;  
 OC Formicarius.  
 OX NCBI\_TaxID=81884;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and

RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY056993; AAL18110.1; -.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109689 MW; 7E8FD7C6AABE69BF CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|:: |||  
Db 662 GILRTFRFVF 671

#### RESULT 74

Q8QG18

ID Q8QG18 PRELIMINARY; PRT; 957 AA.  
AC Q8QG18;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Mimus patagonicus.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Mimidae; Mimus.  
OX NCBI\_TaxID=175008;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057005; AAL18122.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.

FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109588 MW; BE1867653B3846BF CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 662 GILRTFRFVF 671

# RESULT 75

## Q8QG03

ID Q8QG03 PRELIMINARY; PRT; 957 AA.  
AC Q8QG03;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Pipra coronata.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Pipridae; Pipra.  
OX NCBI\_TaxID=175011;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057020; AAL18137.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 957 957  
SQ SEQUENCE 957 AA; 109998 MW; 8EEBF14E2CC740F CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 662 GILRTFRFVF 671

RESULT 76

Q8QG11

ID Q8QG11 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG11;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Orthonyx spaldingii (Chowchilla).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea;  
 OC Orthonychidae; Orthonyx.  
 OX NCBI\_TaxID=38397;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057012; AAL18129.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109695 MW; 58E859B27C1C5A35 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

RESULT 77

Q8QFY1

ID Q8QFY1 PRELIMINARY; PRT; 957 AA.  
 AC Q8QFY1;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).

GN RAG-1.  
 OS *Zosterops senegalensis* (African yellow white-eye).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Zosteropidae; Zosterops.  
 OX NCBI\_TaxID=135989;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY057042; AAL18159.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS00889; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109586 MW; D64FF8675690F263 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |::: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 78

Q8QG17

ID Q8QG17 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG17;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS *Monarcha axillaris*.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;  
 OC Dicrurinae; Monarcha.  
 OX NCBI\_TaxID=175126;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and

RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER..  
 DR EMBL; AY057006; AAL18123.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109599 MW; 4D8A5BEFE1FC57C2 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 79

Q8QG45

ID Q8QG45 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG45;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Alauda arvensis.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea; Alaudidae;  
 OC Alauda.  
 OX NCBI\_TaxID=88112;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER..  
 DR EMBL; AY056978; AAL18095.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.



DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 109747 MW; 5CC650DBD4992D14 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;  
 Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 662 GILRTFRFVF 671

# RESULT 80

Q8QG35

ID Q8QG35 PRELIMINARY; PRT; 957 AA.  
 AC Q8QG35;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Coracina lineata.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Corvoidea; Corvidae;  
 OC Corvinae; Coracina.  
 OX NCBI\_TaxID=175001;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056988; AAL18105.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 957 957  
 SQ SEQUENCE 957 AA; 110100 MW; AD79CFD886607167 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 957;

Best Local Similarity 60.0%; Pred. No. 3.9e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 662 GILRTFRFVF 671

RESULT 81

Q8QG07

ID Q8QG07 PRELIMINARY; PRT; 961 AA.  
AC Q8QG07;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Parula americana (Northern parula).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
OC Fringillidae; Emberizinae; Parula.  
OX NCBI\_TaxID=125947;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057016; AAL18133.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 961 961  
SQ SEQUENCE 961 AA; 110299 MW; 8B3CADEFCC94445C CRC64;

Query Match 59.3%; Score 35; DB 13; Length 961;  
Best Local Similarity 60.0%; Pred. No. 4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 666 GILRTFRFVF 675

RESULT 82

Q8QFY8

ID Q8QFY8 PRELIMINARY; PRT; 961 AA.  
AC Q8QFY8;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Thraupis cyanocephala.  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
OC Fringillidae; Emberizinae; Thraupis.  
OX NCBI\_TaxID=175016;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21829499; PubMed=11839199;  
RA Barker F.K., Barrowclough G.F., Groth J.G.;  
RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
RT biogeographic implications of an analysis of nuclear DNA sequence  
RT data.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
DR EMBL; AY057035; AAL18152.1; -.  
DR GO; GO:0005489; F:electron transporter activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001841; Znf\_ring.  
DR Pfam; PF00097; zf-C3HC4; 1.  
DR SMART; SM00184; RING; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
DR PROSITE; PS00518; ZF\_RING\_1; 1.  
DR PROSITE; PS50089; ZF\_RING\_2; 1.  
KW Metal-binding; Zinc; Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 961 961  
SQ SEQUENCE 961 AA; 110234 MW; 1E5C25FF245D38E7 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 961;  
Best Local Similarity 60.0%; Pred. No. 4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: |||  
Db 666 GILRTFRFVF 675

# RESULT 83

Q8QG41  
ID Q8QG41 PRELIMINARY; PRT; 961 AA.  
AC Q8QG41;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG-1.  
OS Cardinalis cardinalis (Northern cardinal).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;

OC Fringillidae; Emberizinae; Cardinalini; Cardinalis.  
 OX NCBI\_TaxID=98964;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056982; AAL18099.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 961 961  
 SQ SEQUENCE 961 AA; 110132 MW; 565971242185C97B CRC64;

Query Match 59.3%; Score 35; DB 13; Length 961;  
 Best Local Similarity 60.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
 |:|:: ||||  
 Db 666 GILRTFRFVF 675

#### RESULT 84

##### Q8QG31

ID Q8QG31 PRELIMINARY; PRT; 961 AA.  
 AC Q8QG31;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Emberiza schoeniclus (reed bunting).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea;  
 OC Fringillidae; Emberizinae; Emberiza.  
 OX NCBI\_TaxID=30425;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21829499; PubMed=11839199;  
 RA Barker F.K., Barrowclough G.F., Groth J.G.;  
 RT "A phylogenetic hypothesis for passerine birds: taxonomic and  
 RT biogeographic implications of an analysis of nuclear DNA sequence  
 RT data.";  
 RL Proc. R. Soc. Lond., B, Biol. Sci. 269:295-308(2002).

CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AY056992; AAL18109.1; -.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR InterPro; IPR000345; CytC\_heme\_BS.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 961 961  
 SQ SEQUENCE 961 AA; 110225 MW; 64B5E5E47BD5DCA9 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 961;  
 Best Local Similarity 60.0%; Pred. No. 4e+02;  
 Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|:: |||  
 Db 666 GILRTFRFVF 675

# RESULT 85

Q9W6P8

ID Q9W6P8 PRELIMINARY; PRT; 961 AA.  
 AC Q9W6P8;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG-1.  
 OS Passer montanus (Tree sparrow).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Passeriformes; Passeroidea; Passeridae;  
 OC Passer.  
 OX NCBI\_TaxID=9160;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99310776; PubMed=10381315;  
 RA Groth J.G., Barrowclough G.F.;  
 RT "Basal divergences in birds and the phylogenetic utility of the  
 RT nuclear RAG-1 gene."  
 RL Mol. Phylogenet. Evol. 12:115-123(1999).  
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.  
 DR EMBL; AF143738; AAD34964.1; -.  
 DR HSSP; P15919; 1RMD.  
 DR InterPro; IPR001841; Znf\_ring.  
 DR Pfam; PF00097; zf-C3HC4; 1.  
 DR SMART; SM00184; RING; 1.  
 DR PROSITE; PS00518; ZF\_RING\_1; 1.  
 DR PROSITE; PS50089; ZF\_RING\_2; 1.  
 KW Metal-binding; Zinc; Zinc-finger.  
 FT NON\_TER 1 1  
 FT NON\_TER 961 961

SQ SEQUENCE 961 AA; 110229 MW; 2CC607DF1AC10580 CRC64;

Query Match 59.3%; Score 35; DB 13; Length 961;  
Best Local Similarity 60.0%; Pred. No. 4e+02;  
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 GVLQSVRFVF 12  
|:|:: ||||  
Db 666 GILRTFRFVF 675

RESULT 86

Q8SZP4

ID Q8SZP4 PRELIMINARY; PRT; 995 AA.  
AC Q8SZP4;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE RE71565p.  
GN CG8414.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkeley;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,  
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,  
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,  
RA Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,  
RA Celniker S.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AY070613; AAL48084.1; -.  
DR FlyBase; FBgn0034073; CG8414.  
SQ SEQUENCE 995 AA; 111109 MW; 4DDB9E6EDA5EEF6B CRC64;

Query Match 59.3%; Score 35; DB 5; Length 995;  
Best Local Similarity 72.7%; Pred. No. 4.1e+02;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
||| || |||  
Db 955 QGVSSVPFVF 965

RESULT 87

Q9V7I8

ID Q9V7I8 PRELIMINARY; PRT; 995 AA.  
AC Q9V7I8;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE CG8414 protein.  
GN CG8414.

OS *Drosophila melanogaster* (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; *Drosophila*.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,

RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,  
 RA Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,  
 RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,  
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,  
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,  
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,  
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AE003808; AAF58067.2; -.  
 DR FlyBase; FBgn0034073; CG8414.  
 SQ SEQUENCE 995 AA; 111207 MW; 36D1904310B27714 CRC64;

Query Match 59.3%; Score 35; DB 5; Length 995;  
 Best Local Similarity 72.7%; Pred. No. 4.1e+02;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 2 QGVLQSVRFVF 12  
 ||| || |||  
 Db 955 QGVSSSVPFVF 965

# RESULT 88

Q88U41

ID Q88U41 PRELIMINARY; PRT; 1249 AA.  
 AC Q88U41;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ATP-dependent nuclease, subunit A.  
 GN REXA OR LP\_2693.  
 OS Lactobacillus plantarum.  
 OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
 OC Lactobacillus.  
 OX NCBI\_TaxID=1590;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NCIMB 8826 / WCFS1;  
 RX MEDLINE=22480296; PubMed=12566566;  
 RA Kleerebezem M., Boekhorst J., van Kranenburg R., Molenaar D.,



RA Kuipers O.P., Leer R., Tarchini R., Peters S.A., Sandbrink H.M.,  
 RA Fiers M.W.E.J., Stiekema W., Klein Lankhorst R.M., Bron P.A.,  
 RA Hoffer S.M., Nierop Groot M.N., Kerkhoven R., De Vries M., Ursing B.,  
 RA De Vos W.M., Siezen R.J.;  
 RT "Complete genome sequence of *Lactobacillus plantarum* WCFS1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:1990-1995(2003).  
 DR EMBL; AL935259; CAD64937.1; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004003; F:ATP dependent DNA helicase activity; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0006281; P:DNA repair; IEA.  
 DR InterPro; IPR000212; UvrD-helicase.  
 DR Pfam; PF00580; UvrD-helicase; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 1249 AA; 141028 MW; 0C270367B7E0ECF3 CRC64;

Query Match 59.3%; Score 35; DB 16; Length 1249;  
 Best Local Similarity 54.5%; Pred. No. 5.2e+02;  
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFV 11  
 |:|:|:|:  
 Db 769 FKGLFQFVRFI 779

# RESULT 89

Q836J8

ID Q836J8 PRELIMINARY; PRT; 1264 AA.  
 AC Q836J8;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Exonuclease REXA.  
 GN REXA OR EF1113.  
 OS *Enterococcus faecalis* (*Streptococcus faecalis*).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 OX NCBI\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=V583 / ATCC 700802;  
 RX MEDLINE=22550857; PubMed=12663927;  
 RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
 RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
 RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,  
 RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,  
 RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,  
 RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;  
 RT "Role of mobile DNA in the evolution of vancomycin-resistant  
 RT *Enterococcus faecalis*.";  
 RL Science 299:2071-2074(2003).  
 DR EMBL; AE016950; AAO80913.1; -.  
 DR TIGR; EF1113; -.  
 DR GO; GO:0005524; F:ATP binding; IEA.  
 DR GO; GO:0004003; F:ATP dependent DNA helicase activity; IEA.  
 DR GO; GO:0003677; F:DNA binding; IEA.  
 DR GO; GO:0004527; F:exonuclease activity; IEA.  
 DR GO; GO:0006281; P:DNA repair; IEA.

DR InterPro; IPR001064; Crystallin.  
DR InterPro; IPR000212; UvrD-helicase.  
DR Pfam; PF00580; UvrD-helicase; 1.  
DR PROSITE; PS00225; CRYSTALLIN\_BETAGAMMA; 1.  
KW Exonuclease; Complete proteome.  
SQ SEQUENCE 1264 AA; 146450 MW; 6DEA770A76C4F5E7 CRC64;

Query Match 59.3%; Score 35; DB 16; Length 1264;  
Best Local Similarity 54.5%; Pred. No. 5.2e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLSVRFV 11  
|:|:| |||:  
Db 766 FRGLFQFVRFI 776

#### RESULT 90

Q7UDW7

ID Q7UDW7 PRELIMINARY; PRT; 151 AA.  
AC Q7UDW7;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN RB11732.  
OS Rhodopirellula baltica.  
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;  
OC Planctomycetaceae; Pirellula.  
OX NCBI\_TaxID=117;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1;  
RX MEDLINE=22735913; PubMed=12835416;  
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,  
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,  
RA Schlesner H., Amann R., Reinhardt R.;  
RT "Complete genome sequence of the marine planctomycete Pirellula sp.  
RT strain 1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).  
DR EMBL; BX294153; CAD79288.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 151 AA; 16870 MW; 5DAE8590A0BBFA5D CRC64;

Query Match 57.6%; Score 34; DB 16; Length 151;  
Best Local Similarity 70.0%; Pred. No. 94;  
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLSVRF 10  
| ||| |:|  
Db 47 FVGVLSLRF 56

#### RESULT 91

Q8ST26

ID Q8ST26 PRELIMINARY; PRT; 204 AA.  
AC Q8ST26;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE HSPC305 10/100.  
 OS Dictyostelium discoideum (Slime mold).  
 OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.  
 OX NCBI\_TaxID=44689;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AX4;  
 RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,  
 RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,  
 RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;  
 RT "Sequence and Analysis of Chromosome 2 of Dictyostelium."  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AC116305; AAL92188.1; -.  
 SQ SEQUENCE 204 AA; 23921 MW; EFCD3733C7BA9C81 CRC64;

Query Match 57.6%; Score 34; DB 5; Length 204;  
 Best Local Similarity 41.7%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12  
 |::|: |:  
 Db 59 FKGILECCSFIF 70

# RESULT 92

O96097

ID O96097 PRELIMINARY; PRT; 207 AA.  
 AC O96097;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE 24 kDa female-specific fat body protein.  
 OS Antheraea yamamai (Japanese oak silkworm).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
 OC Saturniidae; Saturniinae; Saturniini; Antheraea.  
 OX NCBI\_TaxID=7121;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fat body;  
 RA Kajiura Z.;  
 RT "Female-specific fat body protein of the giant silkworm, Antheraea  
 RT yamamai."  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AB022012; BAA36351.1; -.  
 DR InterPro; IPR004045; GST\_Nterm.  
 SQ SEQUENCE 207 AA; 23743 MW; 7E158BF5EEF53F64 CRC64;

Query Match 57.6%; Score 34; DB 5; Length 207;  
 Best Local Similarity 41.7%; Pred. No. 1.3e+02;  
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRVFV 12  
 : || :|:|::|  
 Db 12 YVGVAESIRYLF 23

# RESULT 93

O96096

ID O96096 PRELIMINARY; PRT; 207 AA.  
AC O96096;  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
DE 24 kDa female-specific fat body protein.  
OS *Antheraea pernyi* (Chinese oak silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Saturniidae; Saturniinae; Saturniini; *Antheraea*.  
OX NCBI\_TaxID=7119;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fat body;  
RA Kajiura Z.;  
RT "Female-specific fat body protein of the giant silkworm, *Antheraea pernyi*."  
RL Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AB022011; BAA36350.1; -.  
DR InterPro; IPR004045; GST\_Nterm.  
SQ SEQUENCE 207 AA; 23713 MW; 5CD7009CD2F33B9D CRC64;

Query Match 57.6%; Score 34; DB 5; Length 207;  
Best Local Similarity 41.7%; Pred. No. 1.3e+02;  
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FQGVLSVRFVF 12  
: || :|:|::|  
Db 12 YVGVAESIRYLF 23

# RESULT 94

Q8YTA4

ID Q8YTA4 PRELIMINARY; PRT; 212 AA.  
AC Q8YTA4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein Alr2820.  
GN ALR2820.  
OS *Anabaena* sp. (strain PCC 7120).  
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
OX NCBI\_TaxID=103690;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21595285; PubMed=11759840;  
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
RA Yasuda M., Tabata S.;  
RT "Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena* sp. strain PCC 7120.";

RL DNA Res. 8:205-213(2001).  
DR EMBL; AP003591; BAB74519.1; -.  
DR PIR; AE2158; AE2158.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 212 AA; 23659 MW; AC9803FBE03077B6 CRC64;

Query Match 57.6%; Score 34; DB 16; Length 212;  
Best Local Similarity 63.6%; Pred. No. 1.3e+02;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 QGVLQSVRFVF 12  
||| :| ||:|  
Db 52 QGVYRSDRFLF 62

RESULT 95

Q82ZE3

ID Q82ZE3 PRELIMINARY; PRT; 214 AA.  
AC Q82ZE3;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thiamin pyrophosphokinase family protein.  
GN EF3117.  
OS Enterococcus faecalis (Streptococcus faecalis).  
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
OX NCBI\_TaxID=1351;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=V583 / ATCC 700802;  
RX MEDLINE=22550857; PubMed=12663927;  
RA Paulsen I.T., Banerjee L., Myers G.S.A., Nelson K.E., Seshadri R.,  
RA Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,  
RA Tettelin H., Dodson R.J., Umayam L., Brinkac L., Beanan M.,  
RA Daugherty S., DeBoy R.T., Durkin S., Kolonay J., Madupu R., Nelson W.,  
RA Vamathevan J., Tran B., Upton J., Hansen T., Shetty J., Khouri H.,  
RA Utterback T., Radune D., Ketchum K.A., Dougherty B.A., Fraser C.M.;  
RT "Role of mobile DNA in the evolution of vancomycin-resistant  
RT Enterococcus faecalis.";  
RL Science 299:2071-2074(2003).  
DR EMBL; AE016956; AA082797.1; -.  
DR TIGR; EF3117; -.  
DR GO; GO:0016301; F:kinase activity; IEA.  
DR InterPro; IPR007371; TPK\_catalytic.  
DR Pfam; PF04263; TPK\_catalytic; 1.  
KW Kinase; Complete proteome.  
SQ SEQUENCE 214 AA; 24145 MW; D1B0080AABD805C8 CRC64;

Query Match 57.6%; Score 34; DB 16; Length 214;  
Best Local Similarity 66.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLSVR 9  
|||||:|  
Db 121 FQGVLRQIR 129

RESULT 96

Q9LMB5

ID Q9LMB5 PRELIMINARY; PRT; 223 AA.  
AC Q9LMB5;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE F14D16.28.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S.,  
RA Kim C., Altafi H., Bei Q., Chin C., Chiou J., Choi E., Conn L.,  
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B., Lee J.,  
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,  
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,  
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,  
RA Ecker J.R.;  
RT "Genomic sequence for Arabidopsis thaliana BAC F14D16 from chromosome  
RT I.";  
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Ecker J.R.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,  
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,  
RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,  
RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,  
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,  
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,  
RA Theologis A., Ecker J.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.  
DR EMBL; AC068602; AAF79296.1; -.  
DR GO; GO:0005634; C:nucleus; IEA.  
DR GO; GO:0005732; C:small nucleolar ribonucleoprotein complex; IEA.  
DR GO; GO:0008248; F:pre-mRNA splicing factor activity; IEA.  
DR GO; GO:0006371; P:mRNA splicing; IEA.  
DR InterPro; IPR001163; snRNP\_Sm.  
DR Pfam; PF01423; LSM; 2.  
DR SMART; SM00651; Sm; 1.  
SQ SEQUENCE 223 AA; 25384 MW; EDB4EC4BC11ADC59 CRC64;

Query Match 57.6%; Score 34; DB 10; Length 223;  
Best Local Similarity 58.3%; Pred. No. 1.4e+02;  
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 FQGVLSVRVFV 12  
|: ||::| |||  
Db 114 FEMVLRNVSFVF 125

RESULT 97

Q99NN3

ID Q99NN3 PRELIMINARY; PRT; 243 AA.  
 AC Q99NN3;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG1.  
 OS Dipodomys heermanni (Kangaroo rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Heteromyidae;  
 OC Dipodomysinae; Dipodomys.  
 OX NCBI\_TaxID=10018;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082082; PubMed=11214319;  
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 RA O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals."  
 RL Nature 409:614-618(2001).  
 DR EMBL; AY011888; AAG38406.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 243 243  
 SQ SEQUENCE 243 AA; 27198 MW; 31F8D7A0457B7022 CRC64;

Query Match 57.6%; Score 34; DB 11; Length 243;  
 Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
 |:|:: ||:|  
 Db 104 GILRTFRFIF 113

RESULT 98

Q9BEX1

ID Q9BEX1 PRELIMINARY; PRT; 258 AA.  
 AC Q9BEX1;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE Recombination activating protein 1 (Fragment).  
 GN RAG1.  
 OS Nycteris thebaica (Egyptian slit-faced bat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Nycteridae; Nycteris.  
 OX NCBI\_TaxID=59467;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21082082; PubMed=11214319;  
 RA Murphy W.J., Eizirik E., Johnson W.E., Zhang Y.P., Ryder O.A.,  
 RA O'Brien S.J.;  
 RT "Molecular phylogenetics and the origins of placental mammals."  
 RL Nature 409:614-618(2001).  
 DR EMBL; AY011907; AAG38425.1; -.

FT NON\_TER 1 1  
FT NON\_TER 258 258  
SQ SEQUENCE 258 AA; 29047 MW; E3B15961E961E94F CRC64;

Query Match 57.6%; Score 34; DB 6; Length 258;  
Best Local Similarity 50.0%; Pred. No. 1.6e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|::|||  
Db 109 GILRTFRFIF 118

RESULT 99

Q8WN46

ID Q8WN46 PRELIMINARY; PRT; 262 AA.  
AC Q8WN46;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Recombination activating protein 1 (Fragment).  
GN RAG1.  
OS Nycteris thebaica (Egyptian slit-faced bat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Chiroptera; Microchiroptera; Nycteridae; Nycteris.  
OX NCBI\_TaxID=59467;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Teeling E.C., Madsen O., Van Den Bussche R., deJong W.W.,  
RA Stanhope M.J., Springer M.S.;  
RT "Microbat paraphyly and the convergent evolution of a key innovation  
RT in Old World rhinolophoid microbats."  
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2002).  
DR EMBL; AF447513; AAL50661.1; -.  
FT NON\_TER 1 1  
FT NON\_TER 262 262  
SQ SEQUENCE 262 AA; 30639 MW; C17A2540048808C7 CRC64;

Query Match 57.6%; Score 34; DB 6; Length 262;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 3 GVLQSVRFVF 12  
|:|::|||  
Db 16 GILRTFRFIF 25

RESULT 100

P70972

ID P70972 PRELIMINARY; PRT; 265 AA.  
AC P70972;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE YBAF protein.  
GN YBAF.  
OS Bacillus subtilis.



OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
 RT *subtilis*.";  
 RL Nature 390:249-256(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE OF 179-265 FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97124188; PubMed=8969501;  
 RA Yasumoto K., Liu H., Jeong S.M., Ohashi Y., Kakinuma S., Tanaka K.,  
 RA Kawamura F., Yoshikawa H., Takahashi H.;  
 RT "Sequence analysis of a 50 kb region between *spo0H* and *rrnH* on the  
 RT *Bacillus subtilis* chromosome.";  
 RL Microbiology 142:3039-3046(1996).  
 DR EMBL; Z99104; CAB11923.1; -.  
 DR EMBL; D64126; BAA10986.1; -.  
 DR PIR; F69742; F69742.  
 DR GO; GO:0015087; F:cobalt ion transporter activity; IEA.  
 DR GO; GO:0006824; P:cobalt ion transport; IEA.  
 DR GO; GO:0009236; P:vitamin B12 biosynthesis; IEA.

DR InterPro; IPR003339; CbiQ.  
DR Pfam; PF02361; CbiQ; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 265 AA; 29692 MW; 45A6673B6AA0C57E CRC64;

Query Match 57.6%; Score 34; DB 16; Length 265;  
Best Local Similarity 54.5%; Pred. No. 1.7e+02;  
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 QGVLQSVRFVF 12  
||: |:|:|:  
Db 109 QGIFISLRFVY 119

Search completed: April 7, 2004, 19:07:43  
Job time : 45.2208 secs

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:40:26 ; Search time 9.66234 Seconds  
 (without alignments)  
 48.087 Million cell updates/sec

Title: US-10-030-735-22  
 Perfect score: 41  
 Sequence: 1 XGVLQNVRF 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	40	97.6	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	31	75.6	57	4	US-09-621-976-5916	Sequence 5916, Ap
3	31	75.6	238	6	5405943-2	Patent No. 5405943
4	31	75.6	400	4	US-09-489-039A-11916	Sequence 11916, A
5	31	75.6	406	4	US-09-543-681A-6072	Sequence 6072, Ap
6	31	75.6	436	6	5405943-4	Patent No. 5405943
7	31	75.6	911	2	US-08-928-692-59	Sequence 59, Appl
8	31	75.6	911	4	US-09-339-972-59	Sequence 59, Appl
9	31	75.6	916	2	US-08-928-692-58	Sequence 58, Appl
10	31	75.6	916	4	US-09-339-972-58	Sequence 58, Appl
11	29	70.7	415	3	US-09-198-956-6	Sequence 6, Appli

12	29	70.7	415	4	US-09-670-141-6	Sequence 6, Appli
13	29	70.7	561	4	US-09-252-991A-20870	Sequence 20870, A
14	29	70.7	609	4	US-09-396-149-9	Sequence 9, Appli
15	29	70.7	715	2	US-08-484-993B-10	Sequence 10, Appl
16	29	70.7	715	2	US-08-484-158B-10	Sequence 10, Appl
17	29	70.7	715	2	US-08-484-596A-10	Sequence 10, Appl
18	29	70.7	715	2	US-08-480-150A-10	Sequence 10, Appl
19	29	70.7	715	3	US-08-458-731-10	Sequence 10, Appl
20	29	70.7	715	3	US-08-149-223A-10	Sequence 10, Appl
21	29	70.7	1350	2	US-08-319-866-9	Sequence 9, Appli
22	28	68.3	166	3	US-08-679-006-31	Sequence 31, Appl
23	28	68.3	172	4	US-09-107-532A-4310	Sequence 4310, Ap
24	28	68.3	172	4	US-09-107-532A-6898	Sequence 6898, Ap
25	28	68.3	175	2	US-08-737-980-2	Sequence 2, Appli
26	28	68.3	218	4	US-10-026-045-1	Sequence 1, Appli
27	28	68.3	305	4	US-09-489-039A-9736	Sequence 9736, Ap
28	28	68.3	342	4	US-09-134-001C-4071	Sequence 4071, Ap
29	28	68.3	345	4	US-09-252-991A-28350	Sequence 28350, A
30	28	68.3	399	4	US-09-328-352-7632	Sequence 7632, Ap
31	28	68.3	402	4	US-09-328-352-4281	Sequence 4281, Ap
32	28	68.3	434	4	US-09-408-020-46	Sequence 46, Appl
33	28	68.3	442	4	US-09-107-532A-6254	Sequence 6254, Ap
34	28	68.3	507	4	US-09-540-236-3391	Sequence 3391, Ap
35	28	68.3	550	4	US-09-198-452A-613	Sequence 613, App
36	28	68.3	683	4	US-09-252-991A-26621	Sequence 26621, A
37	28	68.3	688	4	US-09-252-991A-27096	Sequence 27096, A
38	28	68.3	723	4	US-09-976-594-503	Sequence 503, App
39	28	68.3	748	4	US-09-252-991A-32942	Sequence 32942, A
40	28	68.3	1429	1	US-07-642-002-2	Sequence 2, Appli
41	28	68.3	1429	2	US-08-365-486A-13	Sequence 13, Appl
42	28	68.3	1429	2	US-08-319-866-11	Sequence 11, Appl
43	28	68.3	1429	3	US-08-880-342-13	Sequence 13, Appl
44	28	68.3	1429	4	US-09-661-258-2	Sequence 2, Appli
45	28	68.3	1430	2	US-08-705-625-4	Sequence 4, Appli
46	28	68.3	1430	3	US-09-010-998-5	Sequence 5, Appli
47	28	68.3	1430	3	US-09-220-574-4	Sequence 4, Appli
48	28	68.3	1433	2	US-08-365-486A-21	Sequence 21, Appl
49	28	68.3	1433	3	US-09-123-708-4	Sequence 4, Appli
50	28	68.3	1433	3	US-09-123-624-4	Sequence 4, Appli
51	28	68.3	1433	3	US-08-880-342-21	Sequence 21, Appl
52	28	68.3	1434	2	US-08-365-486A-19	Sequence 19, Appl
53	28	68.3	1434	3	US-08-880-342-19	Sequence 19, Appl
54	28	68.3	1434	4	US-09-661-258-1	Sequence 1, Appli
55	28	68.3	1554	2	US-08-705-625-3	Sequence 3, Appli
56	28	68.3	1554	3	US-09-010-998-6	Sequence 6, Appli
57	28	68.3	1554	3	US-09-220-574-3	Sequence 3, Appli
58	27	65.9	65	4	US-09-540-236-3264	Sequence 3264, Ap
59	27	65.9	121	4	US-09-489-039A-11148	Sequence 11148, A
60	27	65.9	210	3	US-09-081-320-22	Sequence 22, Appl
61	27	65.9	210	4	US-09-574-141A-22	Sequence 22, Appl
62	27	65.9	210	4	US-09-707-780-22	Sequence 22, Appl
63	27	65.9	244	1	US-08-248-466B-15	Sequence 15, Appl
64	27	65.9	246	4	US-09-328-352-4691	Sequence 4691, Ap
65	27	65.9	275	4	US-09-134-001C-3732	Sequence 3732, Ap
66	27	65.9	299	4	US-09-648-004-22	Sequence 22, Appl
67	27	65.9	345	3	US-09-027-900-11	Sequence 11, Appl
68	27	65.9	347	2	US-08-164-292B-13	Sequence 13, Appl

69	27	65.9	347	3	US-08-845-623-13	Sequence 13, Appl
70	27	65.9	347	3	US-08-815-927-13	Sequence 13, Appl
71	27	65.9	347	4	US-09-103-330-13	Sequence 13, Appl
72	27	65.9	347	4	US-09-435-242-13	Sequence 13, Appl
73	27	65.9	362	4	US-09-787-583-2	Sequence 2, Appli
74	27	65.9	382	3	US-08-984-618-12	Sequence 12, Appl
75	27	65.9	389	4	US-09-336-643A-27	Sequence 27, Appl
76	27	65.9	414	4	US-09-489-039A-12578	Sequence 12578, A
77	27	65.9	415	4	US-09-489-039A-10457	Sequence 10457, A
78	27	65.9	446	4	US-09-199-637A-267	Sequence 267, App
79	27	65.9	491	4	US-09-543-681A-8202	Sequence 8202, Ap
80	27	65.9	496	4	US-10-002-720-50	Sequence 50, Appl
81	27	65.9	510	4	US-09-489-039A-11778	Sequence 11778, A
82	27	65.9	538	4	US-09-134-001C-3463	Sequence 3463, Ap
83	27	65.9	539	4	US-09-252-991A-19314	Sequence 19314, A
84	27	65.9	547	4	US-10-164-595-72	Sequence 72, Appl
85	27	65.9	560	4	US-09-543-681A-6771	Sequence 6771, Ap
86	27	65.9	573	4	US-09-252-991A-24021	Sequence 24021, A
87	27	65.9	596	4	US-09-134-000C-6209	Sequence 6209, Ap
88	27	65.9	627	4	US-09-540-236-2428	Sequence 2428, Ap
89	27	65.9	677	4	US-09-328-352-7826	Sequence 7826, Ap
90	27	65.9	689	3	US-09-106-194-2	Sequence 2, Appli
91	27	65.9	1052	4	US-09-252-991A-30591	Sequence 30591, A
92	27	65.9	1128	4	US-09-252-991A-31032	Sequence 31032, A
93	27	65.9	1307	1	US-08-395-246C-2	Sequence 2, Appli
94	26	63.4	88	4	US-09-489-039A-11709	Sequence 11709, A
95	26	63.4	119	4	US-08-858-207A-397	Sequence 397, App
96	26	63.4	129	4	US-09-543-681A-7608	Sequence 7608, Ap
97	26	63.4	131	3	US-08-772-440-23	Sequence 23, Appl
98	26	63.4	131	3	US-08-772-440-27	Sequence 27, Appl
99	26	63.4	134	3	US-08-772-440-16	Sequence 16, Appl
100	26	63.4	134	4	US-09-489-039A-8594	Sequence 8594, Ap
101	26	63.4	145	3	US-08-772-440-14	Sequence 14, Appl
102	26	63.4	156	4	US-09-107-532A-6415	Sequence 6415, Ap
103	26	63.4	158	4	US-09-433-248A-4	Sequence 4, Appli
104	26	63.4	164	3	US-09-009-913-11	Sequence 11, Appl
105	26	63.4	167	3	US-08-772-440-21	Sequence 21, Appl
106	26	63.4	168	4	US-09-107-532A-3745	Sequence 3745, Ap
107	26	63.4	175	3	US-08-772-440-15	Sequence 15, Appl
108	26	63.4	208	4	US-09-134-001C-3478	Sequence 3478, Ap
109	26	63.4	209	3	US-08-772-440-4	Sequence 4, Appli
110	26	63.4	236	4	US-09-134-000C-3636	Sequence 3636, Ap
111	26	63.4	240	4	US-09-198-452A-769	Sequence 769, App
112	26	63.4	250	2	US-08-861-269-5	Sequence 5, Appli
113	26	63.4	250	2	US-09-134-596-5	Sequence 5, Appli
114	26	63.4	250	3	US-09-293-273-5	Sequence 5, Appli
115	26	63.4	255	3	US-09-009-913-9	Sequence 9, Appli
116	26	63.4	258	4	US-09-540-236-2125	Sequence 2125, Ap
117	26	63.4	265	3	US-09-009-913-7	Sequence 7, Appli
118	26	63.4	295	1	US-08-317-223-1	Sequence 1, Appli
119	26	63.4	295	3	US-09-059-849A-1	Sequence 1, Appli
120	26	63.4	295	3	US-09-213-632-1	Sequence 1, Appli
121	26	63.4	295	5	PCT-US95-12675-1	Sequence 1, Appli
122	26	63.4	348	4	US-09-328-352-7509	Sequence 7509, Ap
123	26	63.4	358	2	US-08-823-986A-2	Sequence 2, Appli
124	26	63.4	358	3	US-08-784-385-2	Sequence 2, Appli
125	26	63.4	358	5	PCT-US95-10403-2	Sequence 2, Appli

; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-313-288B-20

Query Match 97.6%; Score 40; DB 1; Length 1170;  
 Best Local Similarity 100.0%; Pred. No. 2.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9  
 |||||  
 Db 210 GVLQNVRF 217

RESULT 2

US-09-621-976-5916  
 ; Sequence 5916, Application US/09621976  
 ; Patent No. 6639063  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Dumas Milne Edwards, J.B.  
 ; APPLICANT: Jobert, S.  
 ; APPLICANT: Giordano, J.Y.  
 ; TITLE OF INVENTION: ESTs and Encoded Human Proteins.  
 ; FILE REFERENCE: GENSET.054PR2  
 ; CURRENT APPLICATION NUMBER: US/09/621,976  
 ; CURRENT FILING DATE: 2000-07-21  
 ; NUMBER OF SEQ ID NOS: 19335  
 ; SOFTWARE: Patent.pm  
 ; SEQ ID NO 5916  
 ; LENGTH: 57  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: SIGNAL  
 ; LOCATION: -15...-1  
 ; NAME/KEY: UNSURE  
 ; LOCATION: 5  
 ; OTHER INFORMATION: Xaa = Cys,Asp,Gly,Tyr  
 US-09-621-976-5916

Query Match 75.6%; Score 31; DB 4; Length 57;  
 Best Local Similarity 62.5%; Pred. No. 7.7;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9  
 |::|||  
 Db 12 GIIQNVSF 19

126	26	63.4	365	4	US-09-252-991A-19530	Sequence 19530, A
127	26	63.4	368	4	US-09-134-000C-4546	Sequence 4546, Ap
128	26	63.4	379	4	US-09-252-991A-26357	Sequence 26357, A
129	26	63.4	389	4	US-09-464-035A-7	Sequence 7, Appli
130	26	63.4	424	4	US-09-328-352-6614	Sequence 6614, Ap
131	26	63.4	444	4	US-09-252-991A-24777	Sequence 24777, A
132	26	63.4	445	4	US-09-543-681A-5757	Sequence 5757, Ap
133	26	63.4	449	4	US-09-489-039A-8009	Sequence 8009, Ap
134	26	63.4	452	4	US-09-134-001C-4173	Sequence 4173, Ap
135	26	63.4	459	4	US-09-252-991A-27280	Sequence 27280, A
136	26	63.4	476	4	US-09-489-039A-8120	Sequence 8120, Ap
137	26	63.4	497	1	US-08-252-492-2	Sequence 2, Appli
138	26	63.4	497	2	US-08-727-126-2	Sequence 2, Appli
139	26	63.4	497	3	US-08-942-761-2	Sequence 2, Appli
140	26	63.4	500	1	US-08-117-083-70	Sequence 70, Appl
141	26	63.4	545	4	US-09-604-957-4	Sequence 4, Appli
142	26	63.4	563	4	US-09-270-957-6	Sequence 6, Appli
143	26	63.4	563	4	US-09-270-957-21	Sequence 21, Appl
144	26	63.4	568	4	US-09-252-991A-29237	Sequence 29237, A
145	26	63.4	593	3	US-09-234-393-24	Sequence 24, Appl
146	26	63.4	593	3	US-09-234-393-50	Sequence 50, Appl
147	26	63.4	593	3	US-09-234-393-52	Sequence 52, Appl
148	26	63.4	593	3	US-09-234-393-54	Sequence 54, Appl
149	26	63.4	593	4	US-09-865-171-24	Sequence 24, Appl
150	26	63.4	593	4	US-09-865-171-50	Sequence 50, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and AviHu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

OM protein - protein search, using sw model

Run on: April 7, 2004, 20:55:10 ; Search time 24.8571 Seconds  
 (without alignments)  
 95.091 Million cell updates/sec

Title: US-10-030-735-22  
 Perfect score: 41  
 Sequence: 1 XGVLQNVRF 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Match	Length	DB	ID	Description
No.						



1	40	97.6	240	12	US-10-419-462-40	Sequence 40, Appl
2	40	97.6	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	40	97.6	831	12	US-09-939-853A-97	Sequence 97, Appl
4	40	97.6	831	12	US-09-939-853A-98	Sequence 98, Appl
5	40	97.6	1152	9	US-09-919-603-1	Sequence 1, Appli
6	40	97.6	1170	12	US-10-211-462-38	Sequence 38, Appl
7	40	97.6	1170	12	US-10-231-956A-482	Sequence 482, App
8	40	97.6	1170	12	US-10-419-462-38	Sequence 38, Appl
9	40	97.6	1170	14	US-10-020-141-12	Sequence 12, Appl
10	40	97.6	1170	14	US-10-017-721-2	Sequence 2, Appli
11	40	97.6	1170	14	US-10-021-660-114	Sequence 114, App
12	40	97.6	1170	14	US-10-008-093-2	Sequence 2, Appli
13	40	97.6	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	36	87.8	1457	9	US-09-772-316-2	Sequence 2, Appli
15	36	87.8	1457	15	US-10-600-058-2	Sequence 2, Appli
16	34	82.9	15	15	US-10-285-394-153	Sequence 153, App
17	33	80.5	3500	14	US-10-153-219-2	Sequence 2, Appli
18	33	80.5	3537	14	US-10-153-219-15	Sequence 15, Appl
19	32	78.0	410	15	US-10-369-493-8711	Sequence 8711, Ap
20	32	78.0	3063	14	US-10-177-293-61	Sequence 61, Appl
21	32	78.0	3063	14	US-10-177-293-63	Sequence 63, Appl
22	32	78.0	3063	14	US-10-301-822-26	Sequence 26, Appl
23	31	75.6	64	9	US-09-738-626-3965	Sequence 3965, Ap
24	31	75.6	75	12	US-10-424-599-165575	Sequence 165575,
25	31	75.6	84	12	US-10-424-599-280578	Sequence 280578,
26	31	75.6	106	12	US-10-424-599-212728	Sequence 212728,
27	31	75.6	215	12	US-10-425-114-69170	Sequence 69170, A
28	31	75.6	226	9	US-09-895-913A-362	Sequence 362, App
29	31	75.6	246	12	US-10-425-114-64020	Sequence 64020, A
30	31	75.6	247	9	US-09-815-242-11342	Sequence 11342, A
31	31	75.6	247	9	US-09-815-242-11514	Sequence 11514, A
32	31	75.6	247	12	US-10-282-122A-58737	Sequence 58737, A
33	31	75.6	274	12	US-10-424-599-173348	Sequence 173348,
34	31	75.6	311	12	US-10-282-122A-70382	Sequence 70382, A
35	31	75.6	312	9	US-09-815-242-5267	Sequence 5267, Ap
36	31	75.6	385	12	US-10-282-122A-69206	Sequence 69206, A
37	31	75.6	385	12	US-10-282-122A-73055	Sequence 73055, A
38	31	75.6	385	15	US-10-369-493-21249	Sequence 21249, A
39	31	75.6	387	9	US-09-815-242-10293	Sequence 10293, A
40	31	75.6	387	9	US-09-815-242-13741	Sequence 13741, A
41	31	75.6	387	12	US-10-282-122A-56141	Sequence 56141, A
42	31	75.6	387	12	US-10-282-122A-56677	Sequence 56677, A
43	31	75.6	387	12	US-10-282-122A-59830	Sequence 59830, A
44	31	75.6	387	12	US-10-282-122A-76296	Sequence 76296, A
45	31	75.6	387	12	US-10-282-122A-77759	Sequence 77759, A
46	31	75.6	387	15	US-10-369-493-498	Sequence 498, App
47	31	75.6	387	15	US-10-369-493-23546	Sequence 23546, A
48	31	75.6	392	12	US-10-282-122A-77057	Sequence 77057, A
49	31	75.6	394	12	US-10-282-122A-46368	Sequence 46368, A
50	31	75.6	394	15	US-10-369-493-23302	Sequence 23302, A
51	31	75.6	446	12	US-10-282-122A-77108	Sequence 77108, A
52	31	75.6	652	15	US-10-369-493-16597	Sequence 16597, A
53	31	75.6	658	12	US-10-282-122A-45950	Sequence 45950, A
54	31	75.6	929	15	US-10-436-323-3	Sequence 3, Appli
55	31	75.6	949	14	US-10-148-351-60	Sequence 60, Appl
56	31	75.6	983	15	US-10-369-493-12973	Sequence 12973, A

57	31	75.6	2358	12	US-10-282-122A-45763	Sequence 45763, A
58	30	73.2	14	15	US-10-436-715-442	Sequence 442, App
59	30	73.2	16	15	US-10-436-715-455	Sequence 455, App
60	30	73.2	122	12	US-10-424-599-143617	Sequence 143617,
61	30	73.2	242	12	US-10-425-114-40353	Sequence 40353, A
62	30	73.2	243	12	US-10-424-599-223767	Sequence 223767,
63	30	73.2	315	12	US-10-282-122A-70417	Sequence 70417, A
64	30	73.2	416	12	US-10-282-122A-63727	Sequence 63727, A
65	30	73.2	430	10	US-09-769-736-6	Sequence 6, Appli
66	30	73.2	461	15	US-10-369-493-8924	Sequence 8924, Ap
67	30	73.2	466	12	US-10-424-599-232021	Sequence 232021,
68	30	73.2	558	14	US-10-342-224-110	Sequence 110, App
69	30	73.2	604	15	US-10-094-749-3042	Sequence 3042, Ap
70	30	73.2	634	12	US-10-282-122A-46137	Sequence 46137, A
71	30	73.2	755	14	US-10-225-567A-651	Sequence 651, App
72	30	73.2	812	12	US-10-416-588-2	Sequence 2, Appli
73	30	73.2	879	15	US-10-292-798-454	Sequence 454, App
74	30	73.2	881	9	US-09-982-736-2	Sequence 2, Appli
75	30	73.2	884	15	US-10-041-615-32	Sequence 32, Appl
76	30	73.2	910	15	US-10-041-615-104	Sequence 104, App
77	30	73.2	926	9	US-09-816-685-2	Sequence 2, Appli
78	30	73.2	926	12	US-10-639-708-2	Sequence 2, Appli
79	30	73.2	926	15	US-10-436-715-20	Sequence 20, Appl
80	29	70.7	52	12	US-10-424-599-268809	Sequence 268809,
81	29	70.7	53	12	US-10-424-599-276257	Sequence 276257,
82	29	70.7	68	9	US-09-764-877-1714	Sequence 1714, Ap
83	29	70.7	68	15	US-10-242-515-1714	Sequence 1714, Ap
84	29	70.7	77	12	US-10-424-599-158611	Sequence 158611,
85	29	70.7	87	12	US-10-424-599-254511	Sequence 254511,
86	29	70.7	96	12	US-10-424-599-185817	Sequence 185817,
87	29	70.7	99	12	US-10-424-599-271725	Sequence 271725,
88	29	70.7	116	12	US-10-424-599-154870	Sequence 154870,
89	29	70.7	138	12	US-10-424-599-250255	Sequence 250255,
90	29	70.7	313	15	US-10-369-493-9589	Sequence 9589, Ap
91	29	70.7	331	12	US-10-282-122A-49753	Sequence 49753, A
92	29	70.7	365	15	US-10-421-654-26	Sequence 26, Appl
93	29	70.7	426	12	US-10-282-122A-52492	Sequence 52492, A
94	29	70.7	464	12	US-10-282-122A-42935	Sequence 42935, A
95	29	70.7	465	10	US-09-769-787-164	Sequence 164, App
96	29	70.7	466	12	US-10-424-599-232205	Sequence 232205,
97	29	70.7	469	12	US-10-425-114-45068	Sequence 45068, A
98	29	70.7	479	15	US-10-369-493-17963	Sequence 17963, A
99	29	70.7	498	12	US-10-425-114-49331	Sequence 49331, A
100	29	70.7	564	15	US-10-369-493-11109	Sequence 11109, A
101	29	70.7	609	14	US-10-372-686-9	Sequence 9, Appli
102	29	70.7	609	14	US-10-371-558-9	Sequence 9, Appli
103	29	70.7	609	14	US-10-375-553-9	Sequence 9, Appli
104	29	70.7	632	15	US-10-369-493-2329	Sequence 2329, Ap
105	29	70.7	759	12	US-10-282-122A-77853	Sequence 77853, A
106	29	70.7	777	10	US-09-733-643-15	Sequence 15, Appl
107	29	70.7	895	14	US-10-345-092-2	Sequence 2, Appli
108	29	70.7	895	14	US-10-345-092-5	Sequence 5, Appli
109	29	70.7	1073	14	US-10-156-761-8117	Sequence 8117, Ap
110	29	70.7	1097	12	US-10-282-122A-45717	Sequence 45717, A
111	29	70.7	1367	15	US-10-369-493-16732	Sequence 16732, A
112	29	70.7	3930	12	US-10-282-122A-46817	Sequence 46817, A
113	28	68.3	60	12	US-10-424-599-245062	Sequence 245062,

114	28	68.3	61	12	US-10-424-599-234831	Sequence 234831,
115	28	68.3	78	12	US-10-424-599-161963	Sequence 161963,
116	28	68.3	83	12	US-10-424-599-220065	Sequence 220065,
117	28	68.3	94	12	US-10-425-114-47963	Sequence 47963, A
118	28	68.3	97	12	US-10-425-114-37672	Sequence 37672, A
119	28	68.3	100	12	US-10-424-599-263347	Sequence 263347,
120	28	68.3	124	11	US-09-833-245-1905	Sequence 1905, Ap
121	28	68.3	129	14	US-10-238-075-393	Sequence 393, App
122	28	68.3	133	12	US-10-424-599-265537	Sequence 265537,
123	28	68.3	175	14	US-10-235-994-12	Sequence 12, Appl
124	28	68.3	175	14	US-10-247-671-184	Sequence 184, App
125	28	68.3	177	12	US-10-424-599-240057	Sequence 240057,
126	28	68.3	188	9	US-09-925-300-1176	Sequence 1176, Ap
127	28	68.3	197	12	US-10-282-122A-51966	Sequence 51966, A
128	28	68.3	216	12	US-10-425-114-47379	Sequence 47379, A
129	28	68.3	228	12	US-10-282-122A-49254	Sequence 49254, A
130	28	68.3	232	12	US-10-424-599-191891	Sequence 191891,
131	28	68.3	245	9	US-09-925-302-686	Sequence 686, App
132	28	68.3	318	12	US-10-425-114-67939	Sequence 67939, A
133	28	68.3	329	12	US-10-424-599-177504	Sequence 177504,
134	28	68.3	335	12	US-10-282-122A-44244	Sequence 44244, A
135	28	68.3	337	9	US-09-815-242-5587	Sequence 5587, Ap
136	28	68.3	337	15	US-10-369-493-11195	Sequence 11195, A
137	28	68.3	340	9	US-09-815-242-12425	Sequence 12425, A
138	28	68.3	340	9	US-09-815-242-13138	Sequence 13138, A
139	28	68.3	341	12	US-10-282-122A-71071	Sequence 71071, A
140	28	68.3	345	12	US-10-424-599-235083	Sequence 235083,
141	28	68.3	345	15	US-10-421-654-60	Sequence 60, Appl
142	28	68.3	345	15	US-10-421-654-70	Sequence 70, Appl
143	28	68.3	345	15	US-10-421-654-86	Sequence 86, Appl
144	28	68.3	347	12	US-10-282-122A-48071	Sequence 48071, A
145	28	68.3	358	12	US-10-425-114-69575	Sequence 69575, A
146	28	68.3	361	16	US-10-389-566-1724	Sequence 1724, Ap
147	28	68.3	373	12	US-10-425-114-66135	Sequence 66135, A
148	28	68.3	393	12	US-10-282-122A-47003	Sequence 47003, A
149	28	68.3	394	14	US-10-156-761-8790	Sequence 8790, Ap
150	28	68.3	394	15	US-10-369-493-17436	Sequence 17436, A

#### ALIGNMENTS

##### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 97.6%; Score 40; DB 12; Length 240;  
Best Local Similarity 100.0%; Pred. No. 1.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9  
|||  
Db 192 GVLQNVRF 199

#### RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 97.6%; Score 40; DB 9; Length 466;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 GVLQNVRF 9  
|||  
Db 263 GVLQNVRF 270

#### RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding  
Same

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:36:09 ; Search time 7.83117 Seconds  
 (without alignments)  
 110.548 Million cell updates/sec

Title: US-10-030-735-22  
 Perfect score: 41  
 Sequence: 1 XGVLQNVRF 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
1	40	97.6	229	2	S57957	thrombospondin 1 -	
2	40	97.6	1170	1	TSHUP1	thrombospondin 1 p	
3	40	97.6	1170	2	A40558	thrombospondin 1 p	
4	32	78.0	207	2	S53801	chitin synthase (E	
5	32	78.0	780	2	T50315	hypothetical prote	
6	32	78.0	889	2	JC6015	chitin synthase (E	
7	32	78.0	1413	2	B82877	conserved hypothet	
8	31	75.6	156	2	S60953	iron-sulfur cofact	
9	31	75.6	186	2	B75421	probable pilin, ty	
10	31	75.6	189	2	B45190	chitin synthase (E	
11	31	75.6	195	2	H45189	chitin synthase (E	
12	31	75.6	198	2	A45190	chitin synthase (E	
13	31	75.6	198	2	G45189	chitin synthase (E	

14	31	75.6	247	1	A64590	probable 3-oxoacyl
15	31	75.6	247	2	B71923	3-oxoacyl-[acyl-ca
16	31	75.6	300	2	D81399	malate dehydrogena
17	31	75.6	308	2	JC5468	leukocidin chain 1
18	31	75.6	311	2	C89968	leukotoxin LukE [i
19	31	75.6	387	1	TVECG	phosphoglycerate k
20	31	75.6	387	2	E91103	phosphoglycerate k
21	31	75.6	387	2	H85948	phosphoglycerate k
22	31	75.6	387	2	AB0113	phosphoglycerate k
23	31	75.6	387	2	AD0875	phosphoglycerate k
24	31	75.6	392	2	F82317	phosphoglycerate k
25	31	75.6	394	1	KIBSGM	phosphoglycerate k
26	31	75.6	394	2	C69675	phosphoglycerate k
27	31	75.6	406	2	G02022	tryptophan oxygena
28	31	75.6	446	2	G82299	phosphoglucomutase
29	31	75.6	467	2	D84938	H+-transporting tw
30	31	75.6	747	2	AE2929	two component resp
31	31	75.6	783	2	A98353	probable transcrip
32	31	75.6	911	2	JC6016	chitin synthase (E
33	31	75.6	916	2	JC2315	chitin synthase (E
34	31	75.6	1114	2	T49517	p63 related protei
35	30	73.2	54	2	S35697	leukocidin chain F
36	30	73.2	70	2	F64066	probable outer mem
37	30	73.2	102	2	G84013	hypothetical prote
38	30	73.2	192	2	S77023	hypothetical prote
39	30	73.2	286	2	C49238	gamma-hemolysin co
40	30	73.2	310	2	S68225	synergohymenotropi
41	30	73.2	312	2	T00160	leukocidin chain S
42	30	73.2	312	2	S32211	leucocidin chain S
43	30	73.2	315	2	PC4078	hlgC-like protein
44	30	73.2	315	2	A49234	leucocidin R S com
45	30	73.2	315	2	JN0626	leukocidin chain S
46	30	73.2	315	2	E90043	gamma-hemolysin co
47	30	73.2	338	2	T32257	hypothetical prote
48	30	73.2	416	2	S72781	probable phosphogl
49	30	73.2	425	2	A96587	hypothetical prote
50	30	73.2	446	2	D89811	hypothetical prote
51	30	73.2	558	2	T48150	stress-induced pro
52	30	73.2	572	2	H86257	protein F5011.2 [i
53	30	73.2	595	2	A43534	Lupus autoantigen
54	30	73.2	914	2	E83901	hypothetical prote
55	30	73.2	1308	2	T05178	hypothetical prote
56	29	70.7	109	2	D82576	hypothetical prote
57	29	70.7	151	2	C57253	tRNA-pseudouridine
58	29	70.7	212	2	A70251	hypothetical prote
59	29	70.7	246	2	H71548	hypothetical prote
60	29	70.7	267	2	S44225	strf protein - Str
61	29	70.7	281	2	S44230	strf protein - Str
62	29	70.7	324	2	C87271	general secretion
63	29	70.7	329	1	JC4251	D-xylose 1-dehydro
64	29	70.7	374	2	S49306	pectase lyase 2 pr
65	29	70.7	411	2	AD2403	hypothetical prote
66	29	70.7	417	2	H83708	hypothetical prote
67	29	70.7	434	2	D72353	lipopolysaccharide
68	29	70.7	442	2	H91097	hypothetical prote
69	29	70.7	450	2	AD3117	polygalacturonase
70	29	70.7	453	2	A98170	hypothetical prote

71	29	70.7	464	2	D85943	probable proteogly
72	29	70.7	464	2	G65071	hypothetical prote
73	29	70.7	508	2	T40249	hypothetical prote
74	29	70.7	609	2	T40625	single-stranded DN
75	29	70.7	644	2	B97885	transporter, trunc
76	29	70.7	682	2	JC7385	multispecific orga
77	29	70.7	708	2	B81038	TonB-dependent rec
78	29	70.7	715	2	S70397	zona pellucida gly
79	29	70.7	740	2	G95153	neuraminidase, pro
80	29	70.7	759	2	AC0368	probable autotrans
81	29	70.7	777	2	T09056	glucan 1,3-beta-gl
82	29	70.7	1071	2	T43255	tricorn proteinase
83	29	70.7	1247	2	T31331	nitric-oxide synth
84	29	70.7	1350	2	T13254	nitric-oxide synth
85	29	70.7	1467	2	T23950	hypothetical prote
86	29	70.7	3124	2	A40020	collagen alpha 1(X
87	29	70.7	4572	2	S57908	hypothetical 527K
88	28	68.3	89	2	B39529	cadherin-associate
89	28	68.3	132	2	T25694	hypothetical prote
90	28	68.3	152	2	S70181	cheW protein - Rho
91	28	68.3	175	2	S75258	hypothetical prote
92	28	68.3	175	2	I39055	Bcl-2 related - hu
93	28	68.3	191	2	T31903	hypothetical prote
94	28	68.3	197	2	D97061	uncharacterized lo
95	28	68.3	202	2	T50396	conserved hypothet
96	28	68.3	218	2	B47712	myelin/oligodendro
97	28	68.3	247	2	A55717	myelin/oligodendro
98	28	68.3	298	2	C87403	FdhD protein [impo
99	28	68.3	334	2	A83225	binding protein co
100	28	68.3	335	2	E89819	hypothetical prote
101	28	68.3	338	2	S56333	carbon-phosphorus
102	28	68.3	338	2	G91264	hypothetical prote
103	28	68.3	338	2	D86105	hypothetical prote
104	28	68.3	357	2	F70433	GcpE protein - Aqu
105	28	68.3	364	2	AF3363	membrane fusion pr
106	28	68.3	376	2	B75260	conserved hypothet
107	28	68.3	379	2	AH2224	hypothetical prote
108	28	68.3	389	2	S76490	hypothetical prote
109	28	68.3	390	1	TVTWG	phosphoglycerate k
110	28	68.3	393	2	H70106	phosphoglycerate k
111	28	68.3	394	2	JQ1399	phosphoglycerate k
112	28	68.3	394	2	G84094	phosphoglycerate k
113	28	68.3	396	2	G87651	phosphoglycerate k
114	28	68.3	399	2	S71368	phosphoglycerate k
115	28	68.3	400	2	AD2322	phosphoglycerate k
116	28	68.3	401	1	TVWTGY	phosphoglycerate k
117	28	68.3	401	2	H96826	hypothetical prote
118	28	68.3	401	2	T03661	phosphoglycerate k
119	28	68.3	403	2	A71484	probable phosphogl
120	28	68.3	403	2	B43260	phosphoglycerate k
121	28	68.3	407	2	A69393	phosphoglycerate k
122	28	68.3	409	2	PN0007	phosphoglycerate k
123	28	68.3	410	2	S68188	phosphoglycerate k
124	28	68.3	410	2	PN0008	phosphoglycerate k
125	28	68.3	411	2	D75408	phosphoglycerate k
126	28	68.3	411	2	B69006	phosphoglycerate k
127	28	68.3	415	1	KIFFPG	phosphoglycerate k

128	28	68.3	417	1	TVCRGC	phosphoglycerate k
129	28	68.3	417	2	A64380	phosphoglycerate k
130	28	68.3	419	2	G71311	probable phosphogl
131	28	68.3	420	1	TVUTG4	phosphoglycerate k
132	28	68.3	421	1	KIUTGC	phosphoglycerate k
133	28	68.3	425	2	S26623	phosphoglycerate k
134	28	68.3	436	2	E87374	RsaA secretion sys
135	28	68.3	438	2	F82944	GTP-binding protei
136	28	68.3	440	1	KIUTGG	phosphoglycerate k
137	28	68.3	440	1	TVUTGB	phosphoglycerate k
138	28	68.3	448	2	D87263	hypothetical prote
139	28	68.3	454	2	D86793	drug-export protei
140	28	68.3	455	1	TVCRGG	phosphoglycerate k
141	28	68.3	478	2	D96603	probable phosphogl
142	28	68.3	480	1	TVWTGC	phosphoglycerate k
143	28	68.3	481	2	T03660	phosphoglycerate k
144	28	68.3	482	2	T07014	phosphoglycerate k
145	28	68.3	488	1	H64055	IMP dehydrogenase
146	28	68.3	489	2	H81912	probable integral
147	28	68.3	489	2	H81127	hypothetical prote
148	28	68.3	492	2	S32491	testosterone 7alph
149	28	68.3	492	2	A31887	testosterone 7alph
150	28	68.3	492	2	A34272	testosterone 7alph

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 97.6%; Score 40; DB 2; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 0.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9  
 |||||  
 Db 192 GVLQNVRF 199

##### RESULT 2



TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.

A;Reference number: A26155; MUID:87057617; PMID:2430973

A;Accession: A26155

A;Molecule type: mRNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.

A;Reference number: A34274; MUID:89291870; PMID:2544587

A;Accession: A34274

A;Molecule type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.

J. Cell Biol. 108, 729-736, 1989

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>

A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-397 <KOB>

A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276; PMID:3461443

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-374, 'RC' <DIX>

A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:22:56 ; Search time 4.20779 Seconds  
 (without alignments)  
 111.372 Million cell updates/sec

Title: US-10-030-735-22  
 Perfect score: 41  
 Sequence: 1 XGVLQNVRF 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
 score greater than or equal to the score of the result being printed,  
 and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	40	97.6	1170	1	TSP1_BOVIN	Q28178 bos taurus
2	40	97.6	1170	1	TSP1_HUMAN	P07996 homo sapien
3	40	97.6	1170	1	TSP1_MOUSE	P35441 mus musculu
4	40	97.6	1173	1	TSP1_XENLA	P35448 xenopus lae
5	32	78.0	889	1	CHSC_ASPFU	Q92197 aspergillus
6	32	78.0	3063	1	CA1C_HUMAN	Q99715 homo sapien
7	32	78.0	3119	1	CA1C_MOUSE	Q60847 mus musculu
8	31	75.6	194	1	CHS2_AJECA	P30577 ajellomyces
9	31	75.6	198	1	CHS2_RHIAT	P30593 rhinocladie
10	31	75.6	386	1	PGK_ECO57	Q8xd03 escherichia
11	31	75.6	386	1	PGK_ECOLI	P11665 escherichia
12	31	75.6	386	1	PGK_PHOLU	Q8gf87 photorhabdu
13	31	75.6	386	1	PGK_SALTY	Q8xg18 salmonella
14	31	75.6	386	1	PGK_VIBPA	Q87111 vibrio para
15	31	75.6	386	1	PGK_VIBVU	Q8dca0 vibrio vuln
16	31	75.6	387	1	PGK_VIBCH	P96154 vibrio chol
17	31	75.6	387	1	PGK_YERPE	Q8zh3 yersinia pe

18	31	75.6	394	1	PGK_BACME	P24269	bacillus me
19	31	75.6	394	1	PGK_BACSU	P40924	bacillus su
20	31	75.6	394	1	PGK_WIGBR	Q8d2p9	wiggleswort
21	31	75.6	406	1	T23O_HUMAN	P48775	homo sapien
22	31	75.6	467	1	FLII_BUCAI	P57178	buchnera ap
23	31	75.6	885	1	CHS3_EXODE	P30602	exophiala d
24	31	75.6	911	1	CHSG_ASPFU	P54267	aspergillus
25	31	75.6	916	1	CHSB_EMENI	Q00757	emericella
26	30	73.2	70	1	Y414_HAEIN	Q57392	haemophilus
27	30	73.2	315	1	HLGC_STAAU	Q07227	staphylococ
28	30	73.2	315	1	LUKS_STAAU	P31716	staphylococ
29	30	73.2	388	1	PGK_BUCAP	Q8k9b3	buchnera ap
30	30	73.2	393	1	PGK_BUCBP	P59461	buchnera ap
31	30	73.2	416	1	PGK_MYCLE	P46712	mycobacteri
32	30	73.2	608	1	KU70_MOUSE	P23475	mus musculu
33	29	70.7	259	1	TRMB_AZOSE	Q8g9c6	azoarcus sp
34	29	70.7	302	1	CXA9_RAJER	Q92107	raja erinac
35	29	70.7	329	1	XYL1_KLULA	P49378	kluveromyc
36	29	70.7	396	1	PGK_LEPIN	Q8f5h8	leptospira
37	29	70.7	442	1	SSNA_ECOLI	Q46812	escherichia
38	29	70.7	609	1	RFA1_SCHPO	Q92372	schizosacch
39	29	70.7	682	1	S219_RAT	Q9jhi3	rattus norv
40	29	70.7	715	1	ZP2_CANFA	P47983	canis famil
41	29	70.7	1071	1	TRI_THEAC	P96086	thermoplasm
42	29	70.7	1174	1	NOS_RHOPR	Q26240	rhodnius pr
43	29	70.7	1247	1	NOS_ANOST	O61608	anopheles s
44	29	70.7	1349	1	NOS_DROME	Q27571	drosophila
45	29	70.7	3124	1	CA1C_CHICK	P13944	gallus gall
46	28	68.3	152	1	CHEW_RHOSH	Q60251	rhodobacter
47	28	68.3	175	1	BFL1_HUMAN	Q16548	homo sapien
48	28	68.3	212	1	NOS_SQUAC	Q9i9m2	squalus aca
49	28	68.3	245	1	MOG_RAT	Q63345	rattus norv
50	28	68.3	246	1	MOG_MOUSE	Q61885	mus musculu
51	28	68.3	252	1	SURE_RHILO	Q98lc9	rhizobium l
52	28	68.3	335	1	Y286_STAEP	Q8ctu5	staphylococ
53	28	68.3	335	1	Y479_STAAW	Q8nxy9	staphylococ
54	28	68.3	335	1	Y524_STAAM	Q99w79	staphylococ
55	28	68.3	338	1	PHND_ECOLI	P16682	escherichia
56	28	68.3	357	1	ISPG_AQUAE	O67496	aquifex aeo
57	28	68.3	380	1	PEX2_YARLI	Q99155	yarrowia li
58	28	68.3	389	1	PGK_THETH	P09403	thermus the
59	28	68.3	393	1	PGK_BORBU	Q59181	borrelia bu
60	28	68.3	394	1	PGK_BACHD	Q9k714	bacillus ha
61	28	68.3	394	1	PGK_BACST	P18912	bacillus st
62	28	68.3	394	1	PGK_THETN	Q8r965	thermoanaer
63	28	68.3	396	1	PGK_CAUCR	Q9a3f5	caulobacter
64	28	68.3	400	1	PGK_ANASP	Q8ypr1	anabaena sp
65	28	68.3	400	1	PGK_SYNEL	Q8dgp7	synechococc
66	28	68.3	401	1	PGKY_TOBAC	Q42962	nicotiana t
67	28	68.3	401	1	PGKY_WHEAT	P12783	triticum ae
68	28	68.3	403	1	PGK_CHLMU	Q9pln4	chlamydia m
69	28	68.3	403	1	PGK_CHLTR	P94686	chlamydia t
70	28	68.3	405	1	PGK_COREF	Q8ft66	corynebacte
71	28	68.3	405	1	PGK_CORGL	Q01655	corynebacte
72	28	68.3	406	1	PGK_METKA	Q8tuul	methanopyru
73	28	68.3	407	1	PGK_ARCFU	O29119	archaeoglob
74	28	68.3	409	1	PGK_METBR	P20972	methanobact

75	28	68.3	410	1	PGK_METFE	P20971	methanother
76	28	68.3	410	1	PGK_PYRFU	P50316	pyrococcus
77	28	68.3	411	1	PGK_DEIRA	Q9rup2	deinococcus
78	28	68.3	411	1	PGK_METTH	O27121	methanobact
79	28	68.3	415	1	PGK_DROME	Q01604	drosophila
80	28	68.3	416	1	PGK2_METAC	Q8tk32	methanosarc
81	28	68.3	416	1	PGK_METMA	Q8pzk7	methanosarc
82	28	68.3	417	1	PGKB_CRIFA	P08966	crithidia f
83	28	68.3	417	1	PGKB_LEIMA	Q27683	leishmania
84	28	68.3	417	1	PGKB_LEIME	Q27684	leishmania
85	28	68.3	417	1	PGK_METJA	Q58058	methanococc
86	28	68.3	419	1	PGK_TREPA	O83549	treponema p
87	28	68.3	420	1	PGK1_TRYCO	P41760	trypanosoma
88	28	68.3	420	1	PGKE_TRYBB	P08893	trypanosoma
89	28	68.3	421	1	PGKB_TRYBB	P07377	trypanosoma
90	28	68.3	433	1	PGKH_SPIOL	P29409	spinacia ol
91	28	68.3	434	1	GSA_CERSY	O74038	cenarchaeum
92	28	68.3	438	1	TRME_UREPA	Q9prc7	ureaplasma
93	28	68.3	440	1	PGKC_TRYBB	P07378	trypanosoma
94	28	68.3	455	1	PGKC_CRIFA	P08967	crithidia f
95	28	68.3	462	1	PGKH_VOLCA	Q9sbn4	volvox cart
96	28	68.3	478	1	PGKH_ARATH	P50318	arabidopsis
97	28	68.3	479	1	PGKC_LEIMA	P50312	leishmania
98	28	68.3	479	1	PGKC_LEIME	Q27685	leishmania
99	28	68.3	480	1	PGKH_WHEAT	P12782	triticum ae
100	28	68.3	481	1	PGKH_TOBAC	Q42961	nicotiana t
101	28	68.3	487	1	IMDH_PASMU	Q916b7	pasteurella
102	28	68.3	488	1	IMDH_HAEIN	P44334	haemophilus
103	28	68.3	492	1	CPA1_RAT	P11711	rattus norv
104	28	68.3	492	1	CPA2_RAT	P15149	rattus norv
105	28	68.3	492	1	CPAC_MOUSE	P56593	mus musculu
106	28	68.3	505	1	PGKA_TRYBB	P08891	trypanosoma
107	28	68.3	508	1	PGKD_TRYBB	P08892	trypanosoma
108	28	68.3	509	1	PGKG_TRYCO	P41762	trypanosoma
109	28	68.3	529	1	NRD1_SCHPO	Q09702	schizosacch
110	28	68.3	538	1	LEU1_GLOVI	Q7ni93	gloeobacter
111	28	68.3	549	1	Y447_MYCPN	P75130	mycoplasma
112	28	68.3	632	1	KU70_CHICK	O93257	gallus gall
113	28	68.3	737	1	YMA2_YEAST	Q04263	saccharomyc
114	28	68.3	755	1	Y572_CHLPN	Q9z7y1	chlamydia p
115	28	68.3	762	1	ABC9_MOUSE	Q9jj59	mus musculu
116	28	68.3	762	1	ABC9_RAT	Q9qyj4	rattus norv
117	28	68.3	766	1	ABC9_HUMAN	Q9np78	homo sapien
118	28	68.3	842	1	MY1A_RAT	Q62774	rattus norv
119	28	68.3	909	1	LDL1_XENLA	Q99087	xenopus lae
120	28	68.3	909	1	MY1A_MOUSE	O88329	mus musculu
121	28	68.3	1003	1	MYSE_DICDI	Q03479	dictyosteli
122	28	68.3	1006	1	MY1D_HUMAN	O94832	homo sapien
123	28	68.3	1006	1	MY1D_RAT	Q63357	rattus norv
124	28	68.3	1043	1	MY1A_BOVIN	P10568	bos taurus
125	28	68.3	1043	1	MY1A_HUMAN	Q9ubc5	homo sapien
126	28	68.3	1065	1	SED4_YEAST	P25365	saccharomyc
127	28	68.3	1107	1	MY1B_MOUSE	P46735	mus musculu
128	28	68.3	1109	1	MYSD_DICDI	P34109	dictyosteli
129	28	68.3	1111	1	MYSB_DICDI	P34092	dictyosteli
130	28	68.3	1136	1	MY1B_RAT	Q05096	rattus norv
131	28	68.3	1147	1	MYSB_ACACA	P19706	acanthamoeb

132	28	68.3	1168	1	MYSC_ACACA	P10569	acanthamoeb
133	28	68.3	1169	1	C1GB_BACTZ	Q9zaz6	bacillus th
134	28	68.3	1429	1	NOS1_MOUSE	Q9z0j4	mus musculu
135	28	68.3	1429	1	NOS1_RAT	P29476	rattus norv
136	28	68.3	1434	1	NOS1_HUMAN	P29475	homo sapien
137	28	68.3	1435	1	NOS1_RABIT	O19132	oryctolagus
138	27	65.9	118	1	Y309_PASMU	Q9cnw4	pasteurella
139	27	65.9	261	1	HIS6_DEIRA	Q9rwd7	deinococcus
140	27	65.9	261	1	ZNUB_HAEIN	P44691	haemophilus
141	27	65.9	274	1	DAPF_HAEIN	P44859	haemophilus
142	27	65.9	275	1	MURI_PSESM	Q888b8	pseudomonas
143	27	65.9	291	1	CU59_DROME	Q9vzh1	drosophila
144	27	65.9	292	1	APAH_XYLFA	Q9pbj4	xylella fas
145	27	65.9	295	1	APAH_XYLFT	Q87c83	xylella fas
146	27	65.9	323	1	Y148_BORPE	O30446	bordetella
147	27	65.9	345	1	AMIE_RHOER	Q01360	rhodococcus
148	27	65.9	358	1	BUK_OCEIH	Q8cxe5	oceanobacil
149	27	65.9	382	1	YPFP_BACSU	P54166	bacillus su
150	27	65.9	391	1	PGK_SHEON	Q8eib1	shewanella

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.

AC Q28178; Q28179;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Thrombospondin 1 precursor.

GN THBS1 OR TSP1 OR TSP-1.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Holstein; TISSUE=Tooth;

RX MEDLINE=98173773; PubMed=9507054;

RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,

RA Inoue H.;

RT "cDNA cloning of bovine thrombospondin 1 and its expression in

RT odontoblasts and predentin.";

RL Biochim. Biophys. Acta 1382:17-22(1998).

RN [2]

RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.

RC TISSUE=Aortic endothelium;

RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and

CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,

CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-

CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis

CC and/or maintenance of dentin and dental pulp.

CC    -!- SUBUNIT: Homotrimer; disulfide-linked.  
 CC    -!- TISSUE SPECIFICITY: Odontoblasts.  
 CC    -!- SIMILARITY: Belongs to the thrombospondin family.  
 CC    -!- SIMILARITY: Contains 1 VWFC domain.  
 CC    -!- SIMILARITY: Contains 3 EGF-like domains.  
 CC    -!- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC    -!- SIMILARITY: Contains 7 TSP type-3 domains.  
 CC    -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC    -----  
 CC    This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC    between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC    the European Bioinformatics Institute. There are no restrictions on its  
 CC    use by non-profit institutions as long as its content is in no way  
 CC    modified and this statement is not removed. Usage by and for commercial  
 CC    entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC    or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC    -----  
 DR    EMBL; AB005287; BAA21115.1; -.  
 DR    EMBL; X87618; CAA60950.1; -.  
 DR    EMBL; X87619; CAA60951.1; -.  
 DR    PIR; S55501; S55501.  
 DR    GlycoSuiteDB; Q28178; -.  
 DR    InterPro; IPR001881; EGF\_Ca.  
 DR    InterPro; IPR006209; EGF\_like.  
 DR    InterPro; IPR006210; IEGF.  
 DR    InterPro; IPR000884; TSP1.  
 DR    InterPro; IPR008085; TSP\_1.  
 DR    InterPro; IPR003367; tsp\_3.  
 DR    InterPro; IPR008859; TSPC.  
 DR    InterPro; IPR003129; TSPN.  
 DR    InterPro; IPR001007; VWF\_C.  
 DR    Pfam; PF00008; EGF; 2.  
 DR    Pfam; PF00090; tsp\_1; 3.  
 DR    Pfam; PF02412; tsp\_3; 13.  
 DR    Pfam; PF05735; TSPC; 1.  
 DR    Pfam; PF02210; TSPN; 1.  
 DR    Pfam; PF00093; vwc; 1.  
 DR    PRINTS; PR01705; TSP1REPEAT.  
 DR    SMART; SM00181; EGF; 3.  
 DR    SMART; SM00209; TSP1; 3.  
 DR    SMART; SM00210; TSPN; 1.  
 DR    SMART; SM00214; VWC; 1.  
 DR    PROSITE; PS00022; EGF\_1; FALSE\_NEG.  
 DR    PROSITE; PS01186; EGF\_2; 1.  
 DR    PROSITE; PS50026; EGF\_3; 2.  
 DR    PROSITE; PS50092; TSP1; 3.  
 DR    PROSITE; PS01208; VWFC\_1; 1.  
 DR    PROSITE; PS50184; VWFC\_2; 1.  
 KW    Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 KW    EGF-like domain; Signal.  
 FT    SIGNAL           1       18           BY SIMILARITY.  
 FT    CHAIN           19      1170          THROMBOSPONDIN 1.  
 FT    DOMAIN          19      232          HEPARIN-BINDING (POTENTIAL).  
 FT    DOMAIN          24      221          TSP N-TERMINAL.  
 FT    DOMAIN          316     373          VWFC.  
 FT    DOMAIN          379     429          TSP TYPE-1 1.  
 FT    DOMAIN          435     490          TSP TYPE-1 2.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:33:49 ; Search time 23.4156 Seconds  
 (without alignments)  
 121.272 Million cell updates/sec

Title: US-10-030-735-22  
 Perfect score: 41  
 Sequence: 1 XGVLQNVRF 9

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp Vertebrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description
-----						

1	40	97.6	229	6	Q28194	Q28194 bos taurus
2	40	97.6	496	13	Q7SY84	Q7sy84 xenopus lae
3	40	97.6	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	40	97.6	1171	11	Q80YQ1	Q80yq1 mus musculu
5	36	87.8	1457	12	Q9DY97	Q9dy97 porcine rep
6	36	87.8	1457	12	Q9WJB3	Q9wjb3 porcine rep
7	36	87.8	1457	12	Q9WBQ4	Q9wbq4 porcine rep
8	36	87.8	1457	12	Q91F53	Q91f53 porcine rep
9	36	87.8	1457	12	Q80KX0	Q80kx0 porcine rep
10	36	87.8	1460	12	Q8QQW9	Q8qqw9 porcine rep
11	36	87.8	1460	12	Q8B911	Q8b911 porcine rep
12	36	87.8	1463	12	Q9YN01	Q9yn01 porcine rep
13	36	87.8	1463	12	Q9ENK5	Q9enk5 porcine rep
14	36	87.8	1463	12	Q9E8M9	Q9e8m9 porcine rep
15	36	87.8	1463	12	Q99AV5	Q99av5 porcine rep
16	36	87.8	1463	12	Q99BU5	Q99bu5 porcine rep
17	36	87.8	1463	12	Q7TF56	Q7tf56 porcine rep
18	36	87.8	3956	12	Q9DLN9	Q9dln9 porcine rep
19	36	87.8	3960	12	Q9DLP1	Q9dlp1 porcine rep
20	36	87.8	3960	12	Q9DLN8	Q9dln8 porcine rep
21	36	87.8	3960	12	Q9DLP0	Q9dlp0 porcine rep
22	35	85.4	1034	5	Q7Z2B9	Q7z2b9 trypanosoma
23	33	80.5	595	5	Q86NR6	Q86nr6 drosophila
24	33	80.5	721	5	Q9VTH0	Q9vth0 drosophila
25	33	80.5	793	16	Q89HG9	Q89hg9 bradyrhizob
26	33	80.5	3242	9	Q859P9	Q859p9 bacterioph
27	32	78.0	207	3	Q9URD7	Q9urd7 aspergillus
28	32	78.0	682	6	Q7YQK2	Q7yqk2 bos taurus
29	32	78.0	725	6	Q7YQK1	Q7yqk1 bos taurus
30	32	78.0	780	3	Q9P7W8	Q9p7w8 schizosacch
31	32	78.0	892	3	Q96VF8	Q96vf8 tuber borch
32	32	78.0	893	3	Q8TGD5	Q8tgd5 aspergillus
33	32	78.0	1413	16	Q9PPU7	Q9ppu7 ureaplasma
34	32	78.0	1463	12	Q9J7C0	Q9j7c0 porcine rep
35	32	78.0	4138	5	Q8I1Y3	Q8ily3 plasmodium
36	31	75.6	64	16	Q8NT94	Q8nt94 corynebacte
37	31	75.6	156	3	Q12056	Q12056 saccharomyc
38	31	75.6	186	16	Q9RUZ7	Q9ruz7 deinococcus
39	31	75.6	195	3	P87045	P87045 metarhizium
40	31	75.6	207	3	Q01164	Q01164 magnaporthe
41	31	75.6	232	5	Q8ISI2	Q8isi2 oxytricha l
42	31	75.6	244	5	Q8IBP0	Q8ibp0 plasmodium
43	31	75.6	247	16	Q25286	Q25286 helicobacte
44	31	75.6	247	16	Q9ZLS0	Q9zls0 helicobacte
45	31	75.6	261	16	Q9CP25	Q9cp25 pasteurella
46	31	75.6	300	16	Q9PHY2	Q9phy2 campylobact
47	31	75.6	308	2	Q53703	Q53703 staphylococ
48	31	75.6	308	2	Q53731	Q53731 staphylococ
49	31	75.6	308	9	Q9MBN3	Q9mbn3 staphylococ
50	31	75.6	311	2	Q93UU9	Q93uu9 staphylococ
51	31	75.6	311	16	Q99T53	Q99t53 staphylococ
52	31	75.6	314	2	Q54081	Q54081 staphylococ
53	31	75.6	341	3	Q8X0C9	Q8x0c9 neurospora
54	31	75.6	394	16	Q81X75	Q81x75 bacillus an
55	31	75.6	446	16	Q9KU84	Q9ku84 vibrio chol
56	31	75.6	446	16	Q8DBW4	Q8dbw4 vibrio vuln
57	31	75.6	446	16	Q87LZ7	Q87lz7 vibrio para



58	31	75.6	681	10	Q84YI2	Q84yi2 hordeum vul
59	31	75.6	690	3	Q9C495	Q9c495 arthroderma
60	31	75.6	783	16	Q8UBI1	Q8ubi1 agrobacteri
61	31	75.6	903	3	Q9HFT1	Q9hft1 coccidioide
62	31	75.6	911	3	Q8TG14	Q8tg14 botrytis ci
63	31	75.6	912	3	Q8TFN5	Q8tfn5 colletotric
64	31	75.6	915	3	Q01749	Q01749 penicillium
65	31	75.6	916	3	Q9C164	Q9c164 aspergillus
66	31	75.6	969	2	Q9F424	Q9f424 leuconostoc
67	31	75.6	1566	16	Q81AL7	Q81al7 bacillus ce
68	31	75.6	2358	16	Q81YE8	Q81ye8 bacillus an
69	31	75.6	2454	3	Q9UV56	Q9uv56 emericella
70	31	75.6	2454	3	Q9UVP2	Q9uvp2 emericella
71	30	73.2	102	16	Q9K8U1	Q9k8u1 bacillus ha
72	30	73.2	175	10	Q84RF8	Q84rf8 gossypium b
73	30	73.2	192	16	Q55947	Q55947 synechocyst
74	30	73.2	237	11	Q8K1Y4	Q8kly4 mus musculu
75	30	73.2	268	16	Q87PY7	Q87py7 vibrio para
76	30	73.2	273	16	Q8DG37	Q8dg37 vibrio vuln
77	30	73.2	280	10	Q94BG4	Q94bg4 froelichia
78	30	73.2	286	2	Q9AFB0	Q9afb0 staphylococ
79	30	73.2	286	2	Q9AFA9	Q9afa9 staphylococ
80	30	73.2	303	16	Q8PIH5	Q8pih5 xanthomonas
81	30	73.2	310	2	Q54326	Q54326 staphylococ
82	30	73.2	312	2	O50603	O50603 staphylococ
83	30	73.2	312	2	Q53746	Q53746 staphylococ
84	30	73.2	312	9	O80066	O80066 staphylococ
85	30	73.2	312	16	Q8NWL8	Q8nwl8 staphylococ
86	30	73.2	315	2	Q53691	Q53691 staphylococ
87	30	73.2	315	2	Q53701	Q53701 staphylococ
88	30	73.2	315	16	Q99RL1	Q99rl1 staphylococ
89	30	73.2	338	5	O17054	O17054 caenorhabdi
90	30	73.2	378	8	Q9GDZ5	Q9gdz5 aeschynomen
91	30	73.2	391	2	Q8RQY4	Q8rqy4 cytophaga s
92	30	73.2	391	2	Q8RQW3	Q8rqw3 cytophaga s
93	30	73.2	391	2	Q8RQY5	Q8rqy5 cytophaga s
94	30	73.2	391	2	Q8RQY3	Q8rqy3 cfb-group b
95	30	73.2	400	16	Q7VRG3	Q7vrg3 candidatus
96	30	73.2	413	16	Q98PC8	Q98pc8 rhizobium 1
97	30	73.2	425	10	Q9SLI3	Q9sli3 arabidopsis
98	30	73.2	446	16	Q99WE4	Q99we4 staphylococ
99	30	73.2	446	16	Q8NY23	Q8ny23 staphylococ
100	30	73.2	464	8	Q9BAA8	Q9baa8 melichrus p
101	30	73.2	470	8	O78404	O78404 pentachondr
102	30	73.2	472	10	Q943N5	Q943n5 oryza sativ
103	30	73.2	474	2	Q9F1Y9	Q9fly9 tenacibacul
104	30	73.2	474	2	Q9ETS8	Q9ets8 tenacibacul
105	30	73.2	474	2	Q9F205	Q9f205 tenacibacul
106	30	73.2	474	2	Q9F1Z0	Q9flz0 tenacibacul
107	30	73.2	474	2	Q9F1Y8	Q9fly8 tenacibacul
108	30	73.2	474	2	Q9FAU7	Q9fau7 polaribacte
109	30	73.2	474	2	Q9EU73	Q9eu73 tenacibacul
110	30	73.2	474	2	Q845X1	Q845x1 polaribacte
111	30	73.2	474	8	Q9BAB1	Q9bab1 cyathodes g
112	30	73.2	478	8	Q9BAA1	Q9baa1 styphelia v
113	30	73.2	479	2	Q9LCK3	Q9lck3 pedobacter
114	30	73.2	481	5	Q961I8	Q961i8 drosophila

115	30	73.2	511	8	Q9GI81	Q9gi81 nissolia sc
116	30	73.2	511	8	Q9GI80	Q9gi80 nissolia hi
117	30	73.2	520	10	Q7XLX8	Q7xlx8 oryza sativ
118	30	73.2	521	4	Q9C0I6	Q9c0i6 homo sapien
119	30	73.2	530	10	Q8L724	Q8l724 arabidopsis
120	30	73.2	558	10	Q9STH1	Q9sth1 arabidopsis
121	30	73.2	572	10	Q9LNB6	Q9lnb6 arabidopsis
122	30	73.2	581	16	Q8E6Q5	Q8e6q5 streptococc
123	30	73.2	581	16	Q8E1A1	Q8ela1 streptococc
124	30	73.2	604	4	Q96M94	Q96m94 homo sapien
125	30	73.2	613	11	Q8R124	Q8r124 mus musculu
126	30	73.2	864	17	Q8TJ15	Q8tj15 methanosarc
127	30	73.2	873	10	Q93YI8	Q93yi8 corylus ave
128	30	73.2	879	4	Q8NGU8	Q8ngu8 homo sapien
129	30	73.2	908	11	Q8CG26	Q8cg26 mus musculu
130	30	73.2	914	16	Q9KBB5	Q9kbb5 bacillus ha
131	30	73.2	926	4	Q8NHZ9	Q8nhz9 homo sapien
132	30	73.2	1308	10	O81903	O81903 arabidopsis
133	30	73.2	1411	12	Q06503	Q06503 lactate deh
134	29	70.7	69	12	Q91FK2	Q91fk2 chilo iride
135	29	70.7	101	1	Q8NKU9	Q8nku9 acidianus a
136	29	70.7	109	16	Q9PB67	Q9pb67 xylella fas
137	29	70.7	109	16	Q87BY4	Q87by4 xylella fas
138	29	70.7	138	10	Q8L574	Q8l574 oryza sativ
139	29	70.7	159	5	Q8SUC6	Q8suc6 encephalito
140	29	70.7	212	16	O50803	O50803 borrelia bu
141	29	70.7	246	16	O84163	O84163 chlamydia t
142	29	70.7	247	16	Q81EN4	Q81en4 bacillus ce
143	29	70.7	248	5	Q8SYC6	Q8syc6 drosophila
144	29	70.7	263	17	Q8U360	Q8u360 pyrococcus
145	29	70.7	267	2	Q53815	Q53815 streptomyce
146	29	70.7	280	5	Q9U681	Q9u681 trypanosoma
147	29	70.7	281	2	Q54259	Q54259 streptomyce
148	29	70.7	292	16	Q8E5J2	Q8e5j2 streptococc
149	29	70.7	292	16	Q8DZU5	Q8dzu5 streptococc
150	29	70.7	292	16	Q82V26	Q82v26 nitrosomona

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.  
AC Q28194;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thrombospondin-1 (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 97.6%; Score 40; DB 6; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 1.3;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GVLQNVRF 9  
 |||||  
 Db 192 GVLQNVRF 199

# RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:40:26 ; Search time 12.8831 Seconds  
 (without alignments)  
 48.087 Million cell updates/sec

Title: US-10-030-735-23  
 Perfect score: 61  
 Sequence: 1 FQGV LQN VKFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	58	95.1	1170	1 US-08-313-288B-20	Sequence 20, Appl
2	42	68.9	1172	1 US-08-313-288B-19	Sequence 19, Appl
3	38	62.3	440	1 US-08-307-499-15	Sequence 15, Appl
4	38	62.3	440	3 US-09-299-268-15	Sequence 15, Appl
5	37	60.7	1288	3 US-08-762-428A-6	Sequence 6, Appli
6	36	59.0	40	1 US-07-868-353A-3	Sequence 3, Appli
7	36	59.0	40	2 US-08-407-804-3	Sequence 3, Appli
8	36	59.0	40	3 US-09-124-807-3	Sequence 3, Appli
9	36	59.0	350	1 US-07-868-353A-14	Sequence 14, Appl
10	36	59.0	350	2 US-08-407-804-23	Sequence 23, Appl
11	36	59.0	350	3 US-09-124-807-23	Sequence 23, Appl

12	36	59.0	354	1	US-07-868-353A-12	Sequence 12, Appl
13	36	59.0	354	1	US-07-868-353A-13	Sequence 13, Appl
14	36	59.0	354	1	US-07-868-353A-15	Sequence 15, Appl
15	36	59.0	354	2	US-08-407-804-21	Sequence 21, Appl
16	36	59.0	354	2	US-08-407-804-22	Sequence 22, Appl
17	36	59.0	354	2	US-08-407-804-24	Sequence 24, Appl
18	36	59.0	354	3	US-09-124-807-21	Sequence 21, Appl
19	36	59.0	354	3	US-09-124-807-22	Sequence 22, Appl
20	36	59.0	354	3	US-09-124-807-24	Sequence 24, Appl
21	36	59.0	538	4	US-09-134-000C-4331	Sequence 4331, Ap
22	35	57.4	102	4	US-09-732-210-1730	Sequence 1730, Ap
23	35	57.4	175	4	US-09-252-991A-18826	Sequence 18826, A
24	35	57.4	740	4	US-09-134-000C-6441	Sequence 6441, Ap
25	35	57.4	810	2	US-08-820-170A-34	Sequence 34, Appl
26	35	57.4	810	3	US-09-055-699-34	Sequence 34, Appl
27	35	57.4	810	3	US-09-273-565-34	Sequence 34, Appl
28	35	57.4	810	4	US-09-565-538-34	Sequence 34, Appl
29	35	57.4	810	4	US-09-661-468-34	Sequence 34, Appl
30	35	57.4	810	4	US-09-976-165-34	Sequence 34, Appl
31	34	55.7	171	4	US-09-328-352-8227	Sequence 8227, Ap
32	34	55.7	275	4	US-09-134-001C-3732	Sequence 3732, Ap
33	33	54.1	390	4	US-09-328-352-4891	Sequence 4891, Ap
34	33	54.1	448	4	US-09-107-532A-6632	Sequence 6632, Ap
35	33	54.1	1142	4	US-09-106-075A-89	Sequence 89, Appl
36	33	54.1	2475	3	US-09-413-814-48	Sequence 48, Appl
37	32	52.5	56	4	US-09-205-258-984	Sequence 984, App
38	32	52.5	57	4	US-09-621-976-5916	Sequence 5916, Ap
39	32	52.5	154	4	US-09-732-210-871	Sequence 871, App
40	32	52.5	158	2	US-08-653-402B-6	Sequence 6, Appli
41	32	52.5	162	4	US-09-732-210-527	Sequence 527, App
42	32	52.5	191	4	US-09-205-258-981	Sequence 981, App
43	32	52.5	292	4	US-09-328-352-4894	Sequence 4894, Ap
44	32	52.5	325	4	US-09-328-352-5100	Sequence 5100, Ap
45	32	52.5	358	4	US-09-252-991A-27541	Sequence 27541, A
46	32	52.5	539	4	US-09-800-170-16	Sequence 16, Appl
47	32	52.5	864	4	US-09-134-000C-6025	Sequence 6025, Ap
48	32	52.5	1019	1	US-08-271-364A-7	Sequence 7, Appli
49	32	52.5	1019	2	US-08-222-715B-26	Sequence 26, Appl
50	31.5	51.6	59	4	US-09-673-395A-423	Sequence 423, App
51	31	50.8	113	4	US-09-134-001C-5539	Sequence 5539, Ap
52	31	50.8	131	4	US-09-252-991A-22819	Sequence 22819, A
53	31	50.8	156	4	US-09-056-556-227	Sequence 227, App
54	31	50.8	156	4	US-09-072-596-222	Sequence 222, App
55	31	50.8	156	4	US-09-072-967-227	Sequence 227, App
56	31	50.8	169	4	US-09-134-000C-3511	Sequence 3511, Ap
57	31	50.8	239	3	US-08-896-933-21	Sequence 21, Appl
58	31	50.8	239	3	US-08-896-933-26	Sequence 26, Appl
59	31	50.8	239	3	US-08-896-933-27	Sequence 27, Appl
60	31	50.8	239	4	US-09-314-235-21	Sequence 21, Appl
61	31	50.8	239	4	US-09-314-235-26	Sequence 26, Appl
62	31	50.8	239	4	US-09-314-235-27	Sequence 27, Appl
63	31	50.8	239	4	US-09-144-776B-10	Sequence 10, Appl
64	31	50.8	240	1	US-07-965-668A-3	Sequence 3, Appli
65	31	50.8	240	2	US-08-950-433-3	Sequence 3, Appli
66	31	50.8	240	3	US-09-186-287-3	Sequence 3, Appli
67	31	50.8	255	1	US-08-446-918A-2	Sequence 2, Appli
68	31	50.8	255	2	US-08-580-806-2	Sequence 2, Appli

69	31	50.8	266	4	US-09-414-276-8	Sequence 8, Appli
70	31	50.8	266	4	US-09-144-776B-8	Sequence 8, Appli
71	31	50.8	266	4	US-09-144-776B-14	Sequence 14, Appl
72	31	50.8	368	4	US-09-819-607-5	Sequence 5, Appli
73	31	50.8	408	4	US-09-328-352-6550	Sequence 6550, Ap
74	31	50.8	432	4	US-09-489-039A-7909	Sequence 7909, Ap
75	31	50.8	445	4	US-09-252-991A-18593	Sequence 18593, A
76	31	50.8	489	1	US-08-434-702-4	Sequence 4, Appli
77	31	50.8	516	4	US-09-291-170A-1	Sequence 1, Appli
78	31	50.8	516	4	US-09-724-884-1	Sequence 1, Appli
79	31	50.8	558	3	US-08-836-567-6	Sequence 6, Appli
80	31	50.8	558	4	US-09-606-304-6	Sequence 6, Appli
81	31	50.8	654	3	US-08-560-005-10	Sequence 10, Appl
82	31	50.8	654	3	US-09-418-540-10	Sequence 10, Appl
83	31	50.8	654	4	US-09-969-528-10	Sequence 10, Appl
84	31	50.8	683	4	US-09-252-991A-26621	Sequence 26621, A
85	31	50.8	715	2	US-08-484-993B-10	Sequence 10, Appl
86	31	50.8	715	2	US-08-484-158B-10	Sequence 10, Appl
87	31	50.8	715	2	US-08-484-596A-10	Sequence 10, Appl
88	31	50.8	715	2	US-08-480-150A-10	Sequence 10, Appl
89	31	50.8	715	3	US-08-458-731-10	Sequence 10, Appl
90	31	50.8	715	3	US-08-149-223A-10	Sequence 10, Appl
91	31	50.8	767	3	US-08-836-567-8	Sequence 8, Appli
92	31	50.8	767	4	US-09-606-304-8	Sequence 8, Appli
93	31	50.8	944	4	US-09-134-000C-5578	Sequence 5578, Ap
94	31	50.8	1148	4	US-09-106-075A-90	Sequence 90, Appl
95	30.5	50.0	147	4	US-09-107-532A-7081	Sequence 7081, Ap
96	30.5	50.0	487	4	US-09-107-532A-5731	Sequence 5731, Ap
97	30	49.2	64	4	US-09-328-352-6751	Sequence 6751, Ap
98	30	49.2	127	4	US-09-614-912-32	Sequence 32, Appl
99	30	49.2	132	3	US-09-073-297-16	Sequence 16, Appl
100	30	49.2	153	2	US-08-387-942C-51	Sequence 51, Appl
101	30	49.2	159	4	US-09-149-476-499	Sequence 499, App
102	30	49.2	163	4	US-09-328-352-5779	Sequence 5779, Ap
103	30	49.2	175	3	US-09-230-637-24	Sequence 24, Appl
104	30	49.2	186	1	US-08-026-758-23	Sequence 23, Appl
105	30	49.2	238	3	US-08-896-933-28	Sequence 28, Appl
106	30	49.2	238	4	US-09-314-235-28	Sequence 28, Appl
107	30	49.2	240	4	US-09-198-452A-769	Sequence 769, App
108	30	49.2	245	4	US-09-328-352-6560	Sequence 6560, Ap
109	30	49.2	300	4	US-09-976-594-7	Sequence 7, Appli
110	30	49.2	310	4	US-09-328-352-7313	Sequence 7313, Ap
111	30	49.2	335	4	US-09-252-991A-24116	Sequence 24116, A
112	30	49.2	346	4	US-09-198-452A-60	Sequence 60, Appl
113	30	49.2	362	4	US-09-787-583-2	Sequence 2, Appli
114	30	49.2	382	3	US-08-984-618-12	Sequence 12, Appl
115	30	49.2	390	3	US-08-650-766-7	Sequence 7, Appli
116	30	49.2	390	3	US-08-922-635-6	Sequence 6, Appli
117	30	49.2	390	4	US-09-389-487-7	Sequence 7, Appli
118	30	49.2	409	4	US-09-489-039A-14217	Sequence 14217, A
119	30	49.2	425	4	US-09-328-352-7930	Sequence 7930, Ap
120	30	49.2	446	4	US-09-199-637A-267	Sequence 267, App
121	30	49.2	541	3	US-08-687-590-28	Sequence 28, Appl
122	30	49.2	541	4	US-09-311-311C-25	Sequence 25, Appl
123	30	49.2	559	4	US-09-364-206-47	Sequence 47, Appl
124	30	49.2	565	4	US-09-134-000C-5782	Sequence 5782, Ap
125	30	49.2	627	4	US-09-540-236-2428	Sequence 2428, Ap

126	30	49.2	646	4	US-09-489-039A-12750	Sequence 12750, A
127	30	49.2	651	3	US-08-650-766-6	Sequence 6, Appli
128	30	49.2	651	3	US-08-922-635-5	Sequence 5, Appli
129	30	49.2	651	4	US-09-389-487-6	Sequence 6, Appli
130	30	49.2	703	4	US-09-328-352-7467	Sequence 7467, Ap
131	30	49.2	710	4	US-09-489-039A-14121	Sequence 14121, A
132	30	49.2	741	4	US-09-107-532A-5001	Sequence 5001, Ap
133	30	49.2	811	4	US-09-614-912-22	Sequence 22, Appl
134	30	49.2	816	2	US-08-820-170A-37	Sequence 37, Appl
135	30	49.2	816	3	US-09-055-699-37	Sequence 37, Appl
136	30	49.2	816	3	US-09-273-565-37	Sequence 37, Appl
137	30	49.2	816	4	US-09-565-538-37	Sequence 37, Appl
138	30	49.2	816	4	US-09-661-468-37	Sequence 37, Appl
139	30	49.2	816	4	US-09-976-165-37	Sequence 37, Appl
140	30	49.2	872	2	US-08-387-942C-5	Sequence 5, Appli
141	30	49.2	940	3	US-09-078-347A-1	Sequence 1, Appli
142	30	49.2	940	4	US-09-651-656-101	Sequence 101, App
143	30	49.2	940	4	US-09-650-855-101	Sequence 101, App
144	30	49.2	941	4	US-09-489-039A-8607	Sequence 8607, Ap
145	30	49.2	994	4	US-09-543-681A-7288	Sequence 7288, Ap
146	30	49.2	1022	1	US-08-271-364A-8	Sequence 8, Appli
147	30	49.2	1022	2	US-08-222-715B-27	Sequence 27, Appl
148	30	49.2	1070	3	US-08-922-635-22	Sequence 22, Appl
149	30	49.2	1076	4	US-09-134-001C-4037	Sequence 4037, Ap
150	30	49.2	1242	4	US-09-107-532A-5241	Sequence 5241, Ap

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-313-288B-20

Query Match 95.1%; Score 58; DB 1; Length 1170;  
 Best Local Similarity 91.7%; Pred. No. 0.011;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGV LQNVKFVF 12  
 |||||:|  
 Db 208 FQGV LQNVRFVF 219

RESULT 2

US-08-313-288B-19  
 ; Sequence 19, Application US/08313288B  
 ; Patent No. 5750502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jessell, Thomas M. and Avi Hu Klar  
 ; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
 ; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/313,288B  
 ; FILING DATE: January 5, 1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 7, 2004, 20:55:10 ; Search time 33.1429 Seconds  
(without alignments)  
95.091 Million cell updates/sec

Title: US-10-030-735-23  
Perfect score: 61  
Sequence: 1 FQGV LQNVKFVF 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 150 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	%	Query				
No.	Score	Match Length DB	ID			Description

1	58	95.1	240	12	US-10-419-462-40	Sequence 40, Appl
2	58	95.1	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	58	95.1	831	12	US-09-939-853A-97	Sequence 97, Appl
4	58	95.1	831	12	US-09-939-853A-98	Sequence 98, Appl
5	58	95.1	1152	9	US-09-919-603-1	Sequence 1, Appli
6	58	95.1	1170	12	US-10-211-462-38	Sequence 38, Appl
7	58	95.1	1170	12	US-10-231-956A-482	Sequence 482, App
8	58	95.1	1170	12	US-10-419-462-38	Sequence 38, Appl
9	58	95.1	1170	14	US-10-020-141-12	Sequence 12, Appl
10	58	95.1	1170	14	US-10-017-721-2	Sequence 2, Appli
11	58	95.1	1170	14	US-10-021-660-114	Sequence 114, App
12	58	95.1	1170	14	US-10-008-093-2	Sequence 2, Appli
13	58	95.1	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	42	68.9	15	15	US-10-285-394-153	Sequence 153, App
15	42	68.9	16	9	US-09-822-682-6	Sequence 6, Appli
16	42	68.9	151	9	US-09-925-299-1251	Sequence 1251, Ap
17	42	68.9	151	10	US-09-925-299-1251	Sequence 1251, Ap
18	42	68.9	1168	9	US-09-919-603-2	Sequence 2, Appli
19	42	68.9	1172	9	US-09-919-770-4	Sequence 4, Appli
20	42	68.9	1172	9	US-09-822-682-2	Sequence 2, Appli
21	42	68.9	1172	10	US-09-866-050A-658	Sequence 658, App
22	42	68.9	1172	14	US-10-060-036-171	Sequence 171, App
23	42	68.9	1172	14	US-10-020-141-14	Sequence 14, Appl
24	42	68.9	1172	14	US-10-017-724-2	Sequence 2, Appli
25	42	68.9	1172	14	US-10-301-822-203	Sequence 203, App
26	42	68.9	1172	14	US-10-008-093-4	Sequence 4, Appli
27	42	68.9	1172	14	US-10-007-781-2	Sequence 2, Appli
28	42	68.9	1172	15	US-10-372-683-16	Sequence 16, Appl
29	38	62.3	465	10	US-09-769-787-164	Sequence 164, App
30	37	60.7	59	11	US-09-864-408A-8466	Sequence 8466, Ap
31	37	60.7	318	12	US-10-425-114-67939	Sequence 67939, A
32	37	60.7	345	12	US-10-424-599-235083	Sequence 235083,
33	37	60.7	426	12	US-10-282-122A-52492	Sequence 52492, A
34	36	59.0	40	10	US-09-789-996-3	Sequence 3, Appli
35	36	59.0	157	10	US-09-952-680A-33	Sequence 33, Appl
36	36	59.0	350	10	US-09-789-996-23	Sequence 23, Appl
37	36	59.0	350	10	US-09-952-680A-24	Sequence 24, Appl
38	36	59.0	350	15	US-10-352-843-15	Sequence 15, Appl
39	36	59.0	354	10	US-09-789-996-21	Sequence 21, Appl
40	36	59.0	354	10	US-09-789-996-22	Sequence 22, Appl
41	36	59.0	354	10	US-09-789-996-24	Sequence 24, Appl
42	36	59.0	354	10	US-09-952-680A-25	Sequence 25, Appl
43	36	59.0	357	9	US-09-984-292-7	Sequence 7, Appli
44	36	59.0	357	9	US-09-989-497-7	Sequence 7, Appli
45	36	59.0	432	12	US-10-424-599-228725	Sequence 228725,
46	36	59.0	939	9	US-09-815-242-10797	Sequence 10797, A
47	36	59.0	939	12	US-10-282-122A-57053	Sequence 57053, A
48	36	59.0	3712	12	US-10-037-417-48	Sequence 48, Appl
49	36	59.0	3712	12	US-10-037-417-51	Sequence 51, Appl
50	36	59.0	3712	13	US-10-108-605-103	Sequence 103, App
51	35	57.4	43	12	US-10-424-599-183491	Sequence 183491,
52	35	57.4	45	12	US-10-424-599-275218	Sequence 275218,
53	35	57.4	48	12	US-10-424-599-197249	Sequence 197249,
54	35	57.4	53	12	US-10-424-599-276257	Sequence 276257,
55	35	57.4	72	12	US-10-424-599-147545	Sequence 147545,
56	35	57.4	106	12	US-10-424-599-242530	Sequence 242530,

57	35	57.4	240	12	US-10-425-114-41336	Sequence 41336, A
58	35	57.4	365	15	US-10-421-654-26	Sequence 26, Appl
59	35	57.4	408	16	US-10-389-566-1540	Sequence 1540, Ap
60	35	57.4	810	9	US-09-976-165-34	Sequence 34, Appl
61	35	57.4	810	14	US-10-342-276-34	Sequence 34, Appl
62	35	57.4	826	12	US-10-424-599-166203	Sequence 166203,
63	35	57.4	1829	15	US-10-435-766-13	Sequence 13, Appl
64	35	57.4	3500	14	US-10-153-219-2	Sequence 2, Appli
65	35	57.4	3537	14	US-10-153-219-15	Sequence 15, Appl
66	34	55.7	51	12	US-10-424-599-198354	Sequence 198354,
67	34	55.7	83	12	US-10-424-599-150515	Sequence 150515,
68	34	55.7	155	12	US-10-424-599-237266	Sequence 237266,
69	34	55.7	274	12	US-10-424-599-173348	Sequence 173348,
70	34	55.7	278	12	US-10-424-599-245776	Sequence 245776,
71	34	55.7	311	12	US-10-282-122A-70382	Sequence 70382, A
72	34	55.7	312	9	US-09-815-242-5267	Sequence 5267, Ap
73	34	55.7	428	15	US-10-421-654-28	Sequence 28, Appl
74	34	55.7	428	15	US-10-421-654-44	Sequence 44, Appl
75	34	55.7	505	15	US-10-369-493-21354	Sequence 21354, A
76	34	55.7	710	12	US-10-282-122A-75098	Sequence 75098, A
77	34	55.7	855	14	US-10-056-790-10	Sequence 10, Appl
78	34	55.7	855	14	US-10-056-790-41	Sequence 41, Appl
79	34	55.7	943	12	US-10-282-122A-73717	Sequence 73717, A
80	34	55.7	954	9	US-09-815-242-13192	Sequence 13192, A
81	34	55.7	954	9	US-09-815-242-13583	Sequence 13583, A
82	34	55.7	1427	12	US-10-282-122A-48514	Sequence 48514, A
83	34	55.7	1457	9	US-09-772-316-2	Sequence 2, Appli
84	34	55.7	1457	15	US-10-600-058-2	Sequence 2, Appli
85	33	54.1	75	12	US-10-424-599-283340	Sequence 283340,
86	33	54.1	87	12	US-10-424-599-254511	Sequence 254511,
87	33	54.1	173	15	US-10-264-237-1799	Sequence 1799, Ap
88	33	54.1	216	15	US-10-264-049-4189	Sequence 4189, Ap
89	33	54.1	226	9	US-09-895-913A-362	Sequence 362, App
90	33	54.1	228	12	US-10-282-122A-49254	Sequence 49254, A
91	33	54.1	247	9	US-09-815-242-11342	Sequence 11342, A
92	33	54.1	247	9	US-09-815-242-11514	Sequence 11514, A
93	33	54.1	247	12	US-10-282-122A-58737	Sequence 58737, A
94	33	54.1	265	15	US-10-094-749-2339	Sequence 2339, Ap
95	33	54.1	279	12	US-10-424-599-261287	Sequence 261287,
96	33	54.1	293	15	US-10-369-493-10330	Sequence 10330, A
97	33	54.1	315	12	US-10-282-122A-70417	Sequence 70417, A
98	33	54.1	318	14	US-10-032-585-7896	Sequence 7896, Ap
99	33	54.1	325	12	US-10-425-114-71821	Sequence 71821, A
100	33	54.1	410	15	US-10-369-493-8711	Sequence 8711, Ap
101	33	54.1	459	13	US-10-108-605-155	Sequence 155, App
102	33	54.1	481	15	US-10-369-493-16410	Sequence 16410, A
103	33	54.1	558	14	US-10-342-224-110	Sequence 110, App
104	33	54.1	1014	12	US-10-424-599-148295	Sequence 148295,
105	33	54.1	1336	15	US-10-369-493-1723	Sequence 1723, Ap
106	33	54.1	1539	14	US-10-072-977-19	Sequence 19, Appl
107	32	52.5	44	12	US-10-424-599-268924	Sequence 268924,
108	32	52.5	46	12	US-10-424-599-144596	Sequence 144596,
109	32	52.5	55	13	US-10-016-634A-119	Sequence 119, App
110	32	52.5	56	10	US-09-933-767-984	Sequence 984, App
111	32	52.5	56	14	US-10-023-282-984	Sequence 984, App
112	32	52.5	92	12	US-10-424-599-241959	Sequence 241959,
113	32	52.5	119	11	US-09-833-245-75	Sequence 75, Appl

114	32	52.5	163	9	US-09-815-242-11071	Sequence 11071, A
115	32	52.5	163	12	US-10-282-122A-58248	Sequence 58248, A
116	32	52.5	163	12	US-10-282-122A-67393	Sequence 67393, A
117	32	52.5	191	10	US-09-933-767-981	Sequence 981, App
118	32	52.5	191	14	US-10-023-282-981	Sequence 981, App
119	32	52.5	203	12	US-10-424-599-234398	Sequence 234398,
120	32	52.5	211	12	US-10-335-977-7519	Sequence 7519, Ap
121	32	52.5	220	12	US-10-424-599-251762	Sequence 251762,
122	32	52.5	224	12	US-10-335-977-7521	Sequence 7521, Ap
123	32	52.5	282	12	US-10-424-599-244115	Sequence 244115,
124	32	52.5	289	15	US-10-369-493-19908	Sequence 19908, A
125	32	52.5	291	12	US-10-424-599-163083	Sequence 163083,
126	32	52.5	305	16	US-10-389-566-594	Sequence 594, App
127	32	52.5	318	14	US-10-259-165-38	Sequence 38, Appl
128	32	52.5	318	14	US-10-259-165-378	Sequence 378, App
129	32	52.5	323	14	US-10-259-165-322	Sequence 322, App
130	32	52.5	359	15	US-10-369-493-10609	Sequence 10609, A
131	32	52.5	370	9	US-09-895-913A-210	Sequence 210, App
132	32	52.5	374	12	US-10-424-599-181131	Sequence 181131,
133	32	52.5	406	12	US-10-335-977-7522	Sequence 7522, Ap
134	32	52.5	409	12	US-10-424-599-234443	Sequence 234443,
135	32	52.5	415	9	US-09-925-301-1360	Sequence 1360, Ap
136	32	52.5	442	12	US-10-282-122A-53275	Sequence 53275, A
137	32	52.5	459	15	US-10-369-493-18328	Sequence 18328, A
138	32	52.5	482	12	US-10-425-114-42018	Sequence 42018, A
139	32	52.5	498	12	US-10-425-114-49331	Sequence 49331, A
140	32	52.5	539	15	US-10-422-536-16	Sequence 16, Appl
141	32	52.5	669	15	US-10-104-047-2373	Sequence 2373, Ap
142	32	52.5	759	12	US-10-282-122A-77853	Sequence 77853, A
143	32	52.5	895	14	US-10-345-092-2	Sequence 2, Appli
144	32	52.5	895	14	US-10-345-092-5	Sequence 5, Appli
145	32	52.5	1300	10	US-09-896-923-3	Sequence 3, Appli
146	32	52.5	1300	15	US-10-223-650-3	Sequence 3, Appli
147	32	52.5	1397	12	US-10-282-122A-50496	Sequence 50496, A
148	32	52.5	1508	15	US-10-369-493-4104	Sequence 4104, Ap
149	32	52.5	1751	15	US-10-435-766-103	Sequence 103, App
150	31	50.8	21	14	US-10-172-425B-17	Sequence 17, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 95.1%; Score 58; DB 12; Length 240;  
Best Local Similarity 91.7%; Pred. No. 0.011;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGV LQN VKFVF 12  
||| ||| ||| : |||  
Db 190 FQGV LQN VRFVF 201

#### RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 95.1%; Score 58; DB 9; Length 466;  
Best Local Similarity 91.7%; Pred. No. 0.023;  
Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQGV LQN VKFVF 12  
||| ||| ||| : |||  
Db 261 FQGV LQN VRFVF 272

#### RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding  
Same

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:36:09 ; Search time 10.4416 Seconds  
 (without alignments)  
 110.548 Million cell updates/sec

Title: US-10-030-735-23  
 Perfect score: 61  
 Sequence: 1 FQGV LQNVKFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	58	95.1	229	2	S57957	thrombospondin 1 -	
2	58	95.1	1170	1	TSHUP1	thrombospondin 1 p	
3	58	95.1	1170	2	A40558	thrombospondin 1 p	
4	42	68.9	1172	1	TSHUP2	thrombospondin 2 p	
5	42	68.9	1172	2	A42587	thrombospondin 2 p	
6	40	65.6	1178	1	A39804	thrombospondin pre	
7	38	62.3	304	2	T34271	hypothetical prote	
8	38	62.3	740	2	G95153	neuraminidase, pro	
9	37	60.7	337	2	E97882	hypothetical prote	
10	37	60.7	431	2	T32359	hypothetical prote	
11	37	60.7	834	2	T39891	probable integral	
12	37	60.7	1288	2	T42756	5-oxoprolinase (AT	
13	36	59.0	212	2	D81929	probable imidazole	

14	36	59.0	350	1	RGHUT1	GTP-binding regula
15	36	59.0	350	1	RGBOT1	GTP-binding regula
16	36	59.0	350	1	RGMST1	GTP-binding regula
17	36	59.0	354	1	RGHUT2	GTP-binding regula
18	36	59.0	354	1	RGBOT2	GTP-binding regula
19	36	59.0	354	2	S24352	gustducin - rat
20	36	59.0	354	2	T20169	hypothetical prote
21	36	59.0	750	1	HYHUN	neprilysin (EC 3.4
22	36	59.0	750	1	HYRTN	neprilysin (EC 3.4
23	36	59.0	751	1	HYRBN	neprilysin (EC 3.4
24	36	59.0	893	2	T36795	probable penicilli
25	36	59.0	1142	2	T39103	probable negative
26	36	59.0	3712	2	S18253	laminin alpha-1 ch
27	35	57.4	82	2	F96625	hypothetical prote
28	35	57.4	102	1	R3PP14	ribosomal protein
29	35	57.4	106	2	C90261	hypothetical prote
30	35	57.4	145	2	T20985	hypothetical prote
31	35	57.4	162	2	B88349	protein F15D4.3 [i
32	35	57.4	199	1	G64070	imidazoleglycerol-
33	35	57.4	439	2	E36840	C14L protein - var
34	35	57.4	439	2	H72154	E10L protein - var
35	35	57.4	439	2	T30787	hypothetical prote
36	35	57.4	439	2	F42507	F10L protein - vac
37	35	57.4	439	2	T28472	hypothetical prote
38	35	57.4	467	2	D84938	H+-transporting tw
39	35	57.4	498	2	B89832	hypothetical prote
40	35	57.4	628	2	AF1108	transporter homolo
41	35	57.4	628	2	AG1469	transporter homolo
42	35	57.4	810	2	T10756	Nel-homolog protei
43	35	57.4	1829	2	AE1864	hypothetical prote
44	35	57.4	1839	2	S77626	mannuronan C-5-epi
45	34.5	56.6	218	2	B42469	polysialic acid tr
46	34.5	56.6	224	2	S12237	kpsT protein - Esc
47	34	55.7	152	2	S44740	C02C2.2 protein -
48	34	55.7	156	2	S60953	iron-sulfur cofact
49	34	55.7	259	2	B46337	sim region ORF2 pr
50	34	55.7	308	2	JC5468	leukocidin chain 1
51	34	55.7	311	2	F70184	ribose/galactose A
52	34	55.7	311	2	C89968	leukotoxin LukE [i
53	34	55.7	385	2	S56224	hypothetical prote
54	34	55.7	425	2	D88115	protein F53C3.11 [
55	34	55.7	457	2	A54604	regulatory protein
56	34	55.7	459	2	B95171	NADH oxidase SP146
57	34	55.7	459	2	B98037	NADH oxidase (EC 1
58	34	55.7	481	2	T15657	hypothetical prote
59	34	55.7	505	2	C69415	carbohydrate kinas
60	34	55.7	515	2	T40649	hypothetical prote
61	34	55.7	905	1	P3XRA4	inner capsid prote
62	34	55.7	943	2	F95021	excinuclease ABC,
63	34	55.7	943	2	C97893	excinuclease ABC c
64	34	55.7	946	2	S71168	Ca2+-transporting
65	34	55.7	957	2	B84099	excinuclease ABC (
66	34	55.7	1020	2	D86402	protein envelope C
67	34	55.7	1020	2	T51925	Ca2+-transporting
68	34	55.7	1020	2	T51926	Ca2+-transporting
69	34	55.7	1038	2	T02634	rep protein homolo
70	34	55.7	1045	2	I51555	recombination acti

71	34	55.7	4767	2	T31345	hypothetical prote
72	33	54.1	54	2	S35697	leukocidin chain F
73	33	54.1	83	2	T12839	hypothetical prote
74	33	54.1	111	2	S52596	probable membrane
75	33	54.1	135	2	S77740	probable phosphotr
76	33	54.1	151	2	C57253	tRNA-pseudouridine
77	33	54.1	186	2	B75421	probable pilin, ty
78	33	54.1	193	2	E86716	acetyl transferase
79	33	54.1	206	2	S70004	hypothetical prote
80	33	54.1	222	2	A26489	placental lactogen
81	33	54.1	247	1	A64590	probable 3-oxoacyl
82	33	54.1	247	2	B71923	3-oxoacyl-[acyl-ca
83	33	54.1	286	2	C49238	gamma-hemolysin co
84	33	54.1	310	2	S68225	synergohymenotropi
85	33	54.1	312	2	T00160	leukocidin chain S
86	33	54.1	312	2	S32211	leucocidin chain S
87	33	54.1	315	2	PC4078	hlgC-like protein
88	33	54.1	315	2	A49234	leucocidin R S com
89	33	54.1	315	2	JN0626	leukocidin chain S
90	33	54.1	315	2	E90043	gamma-hemolysin co
91	33	54.1	371	2	B69451	conserved hypothet
92	33	54.1	407	2	D81313	probable transmemb
93	33	54.1	425	2	A96587	hypothetical prote
94	33	54.1	427	2	A71612	translation releas
95	33	54.1	440	2	F96556	IAA-Ala hydrolase
96	33	54.1	469	2	T33595	hypothetical prote
97	33	54.1	474	2	S07754	NADH2 dehydrogenas
98	33	54.1	513	2	E71683	NADH2 dehydrogenas
99	33	54.1	555	2	T23531	hypothetical prote
100	33	54.1	558	2	T48150	stress-induced pro
101	33	54.1	572	2	H86257	protein F5011.2 [i
102	33	54.1	582	2	T07953	lectin-like protei
103	33	54.1	582	2	T07952	lectin-like protei
104	33	54.1	662	2	H97834	cytochrome c-type
105	33	54.1	670	2	C71630	cytochrome C-type
106	33	54.1	682	2	JC7385	multispecific orga
107	33	54.1	755	2	G90095	hypothetical prote
108	33	54.1	780	2	T50315	hypothetical prote
109	33	54.1	957	2	F69729	excinuclease ABC c
110	33	54.1	1014	2	T04721	Ca2+-transporting
111	33	54.1	1015	2	H84618	probable Ca2+-ATPa
112	33	54.1	1019	2	JC7538	neuronal different
113	33	54.1	1041	2	S42509	Rag-1 protein - ch
114	33	54.1	1042	2	S42511	RAG-1 protein - ra
115	33	54.1	1043	2	A33754	recombination-acti
116	33	54.1	1054	2	H69377	reverse gyrase (to
117	33	54.1	1069	2	C85349	Ca2+-transporting
118	33	54.1	1093	2	T08551	Ca2+-transporting
119	33	54.1	1134	2	S53955	hypothetical prote
120	33	54.1	1142	1	GNVUPH	M polyprotein prec
121	33	54.1	1148	1	JQ1604	M polyprotein prec
122	33	54.1	1336	2	S41794	SEC3 protein - yea
123	33	54.1	1545	2	T42751	sulfonylurea recep
124	33	54.1	1545	2	T46645	sulfonylurea recep
125	33	54.1	2139	2	A35672	crumbs protein - f
126	33	54.1	4572	2	S57908	hypothetical 527K
127	33	54.1	6658	2	T13931	projectin - fruit



128	32	52.5	89	2	B39529	cadherin-associate
129	32	52.5	99	2	AH1886	hypothetical prote
130	32	52.5	102	2	G84013	hypothetical prote
131	32	52.5	110	2	A24444	hypothetical prote
132	32	52.5	122	2	T28199	hypothetical prote
133	32	52.5	138	2	G97814	hypothetical prote
134	32	52.5	144	2	F71215	hypothetical prote
135	32	52.5	154	2	D64359	ribosomal protein
136	32	52.5	163	2	E64083	ribosomal protein
137	32	52.5	224	2	T51875	hypothetical prote
138	32	52.5	241	2	A75065	hypothetical prote
139	32	52.5	241	2	D71167	hypothetical prote
140	32	52.5	298	2	T12084	hypothetical prote
141	32	52.5	316	2	S69659	hypothetical prote
142	32	52.5	317	2	G70313	lipopolysaccharide
143	32	52.5	326	2	S54267	repA protein - Bac
144	32	52.5	331	2	S54263	rep A protein - Ba
145	32	52.5	332	2	I67791	cytochrome P450 2B
146	32	52.5	335	2	T43627	hypothetical prote
147	32	52.5	345	2	A40990	GTP-binding regula
148	32	52.5	349	2	G83605	probable ATP-bindi
149	32	52.5	351	2	T26591	hypothetical prote
150	32	52.5	372	2	S32694	Wnt-1 protein - Ca

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 95.1%; Score 58; DB 2; Length 229;

Best Local Similarity 91.7%; Pred. No. 0.0011;

Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FQGV	LQNV	KFVF	12
Db	190	FQGV	LQNV	RFVF	201

##### RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.

A;Reference number: A26155; MUID:87057617; PMID:2430973

A;Accession: A26155

A;Molecule type: mRNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.

A;Reference number: A34274; MUID:89291870; PMID:2544587

A;Accession: A34274

A;Molecule type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.

J. Cell Biol. 108, 729-736, 1989

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>

A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-397 <KOB>

A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276; PMID:3461443

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-374, 'RC' <DIX>

A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:22:56 ; Search time 5.61039 Seconds  
 (without alignments)  
 111.372 Million cell updates/sec

Title: US-10-030-735-23  
 Perfect score: 61  
 Sequence: 1 FQGV LQN VKFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	58	95.1	1170	1 TSP1_BOVIN	Q28178 bos taurus
2	58	95.1	1170	1 TSP1_HUMAN	P07996 homo sapien
3	58	95.1	1170	1 TSP1_MOUSE	P35441 mus musculu
4	58	95.1	1173	1 TSP1_XENLA	P35448 xenopus lae
5	42	68.9	1172	1 TSP2_HUMAN	P35442 homo sapien
6	42	68.9	1172	1 TSP2_MOUSE	Q03350 mus musculu
7	41	67.2	1170	1 TSP2_BOVIN	Q95116 bos taurus
8	40	65.6	1178	1 TSP2_CHICK	P35440 gallus gall
9	38	62.3	440	1 KRF1_SPVKA	P32216 swinepox vi
10	37	60.7	1288	1 OPLA_MOUSE	Q8k010 mus musculu
11	37	60.7	1288	1 OPLA_RAT	P97608 rattus norv
12	36	59.0	212	1 HIS5_NEIMA	Q9jvh3 neisseria m
13	36	59.0	349	1 GBT1_BOVIN	P04695 bos taurus
14	36	59.0	349	1 GBT1_CANFA	Q28300 canis famil
15	36	59.0	349	1 GBT1_HUMAN	P11488 homo sapien
16	36	59.0	349	1 GBT1_MOUSE	P20612 mus musculu
17	36	59.0	353	1 GBT2_BOVIN	P04696 bos taurus

18	36	59.0	353	1	GBT2_HUMAN	P19087	homo sapien
19	36	59.0	353	1	GBT2_MOUSE	P50149	mus musculu
20	36	59.0	353	1	GBT3_RAT	P29348	rattus norv
21	36	59.0	749	1	NEP_HUMAN	P08473	homo sapien
22	36	59.0	749	1	NEP_MOUSE	Q61391	mus musculu
23	36	59.0	749	1	NEP_RABIT	P08049	oryctolagus
24	36	59.0	749	1	NEP_RAT	P07861	rattus norv
25	36	59.0	3712	1	LMA_DROME	Q00174	drosophila
26	35	57.4	102	1	RT14_PARTE	P15759	paramecium
27	35	57.4	199	1	HIS5_HAEIN	P44340	haemophilus
28	35	57.4	380	1	PEX2_YARLI	Q99155	yarrowia li
29	35	57.4	405	1	KRF1_VACCP	P29884	vaccinia vi
30	35	57.4	439	1	KRF1_VACCC	P21095	vaccinia vi
31	35	57.4	439	1	KRF1_VARV	P33801	variola vir
32	35	57.4	467	1	FLII_BUCAI	P57178	buchnera ap
33	35	57.4	810	1	NEL1_HUMAN	Q92832	homo sapien
34	35	57.4	810	1	NEL1_RAT	Q62919	rattus norv
35	35	57.4	1391	1	RPC1_HUMAN	O14802	homo sapien
36	35	57.4	1829	1	DPOL_THEST	O33845	thermococcu
37	35	57.4	1839	1	ALE3_AZOVI	Q44496	azotobacter
38	34.5	56.6	219	1	KST1_ECOLI	P23888	escherichia
39	34.5	56.6	224	1	KST5_ECOLI	P24586	escherichia
40	34	55.7	243	1	HIS4_HELHP	Q7vhy5	helicobacte
41	34	55.7	250	1	UBIE_COXBU	Q83a90	coxiella bu
42	34	55.7	312	1	OTCC_MYCCC	P59779	mycoplasma
43	34	55.7	385	1	YFD0_YEAST	P43567	saccharomyc
44	34	55.7	457	1	GAL8_KLULA	Q06433	kluveromyc
45	34	55.7	905	1	VP3_AHSV4	P32509	african hor
46	34	55.7	905	1	VP3_AHSV6	O71025	african hor
47	34	55.7	943	1	UVRA_STRPN	Q97sx7	streptococc
48	34	55.7	957	1	UVRA_BACHD	Q9k6y0	bacillus ha
49	34	55.7	1020	1	ACA1_ARATH	Q37145	arabidopsis
50	34	55.7	1045	1	RAG1_XENLA	Q91829	xenopus lae
51	33	54.1	222	1	PLL2_MOUSE	P09586	mus musculu
52	33	54.1	315	1	HLGC_STAAU	Q07227	staphylococ
53	33	54.1	315	1	LUKS_STAAU	P31716	staphylococ
54	33	54.1	349	1	GBT_XENLA	P38407	xenopus lae
55	33	54.1	352	1	IDI2_PYRAE	Q8zyf6	pyrobaculum
56	33	54.1	438	1	ERF1_DROME	Q9vph7	drosophila
57	33	54.1	474	1	NU4M_PARTE	P15581	paramecium
58	33	54.1	682	1	S219_RAT	Q9jhi3	rattus norv
59	33	54.1	957	1	UVRA_BACSU	O34863	bacillus su
60	33	54.1	958	1	UVRA_OCEIH	Q8enj6	oceanobacil
61	33	54.1	1014	1	ACA2_ARATH	O81108	arabidopsis
62	33	54.1	1015	1	ACA7_ARATH	O64806	arabidopsis
63	33	54.1	1041	1	RAG1_CHICK	P24271	gallus gall
64	33	54.1	1042	1	RAG1_RABIT	P34088	oryctolagus
65	33	54.1	1043	1	RAG1_HUMAN	P15918	homo sapien
66	33	54.1	1069	1	ACAA_ARATH	Q9sizr1	arabidopsis
67	33	54.1	1134	1	YML7_YEAST	Q03735	saccharomyc
68	33	54.1	1142	1	VGLM_PHV	P27315	prospect hi
69	33	54.1	1148	1	VGLM_PUUMS	P27312	puumala vir
70	33	54.1	1205	1	PDS5_SCHPO	Q9hff5	schizosacch
71	33	54.1	1336	1	SEC3_YEAST	P33332	saccharomyc
72	33	54.1	1545	1	ACC9_RAT	Q63563	rattus norv
73	33	54.1	1549	1	ACC9_RABIT	P82451	oryctolagus
74	33	54.1	2139	1	CRB_DROME	P10040	drosophila

75	32	52.5	110	1	YCX1_CHLPY	P05720	chlorella p
76	32	52.5	126	1	Y334_BUCBP	Q89ag0	buchnera ap
77	32	52.5	154	1	RL30_METJA	P54046	methanococc
78	32	52.5	162	1	RL10_HAEIN	P44350	haemophilus
79	32	52.5	163	1	RL10_PASMU	Q9ck89	pasteurella
80	32	52.5	188	1	MAUE_METFL	Q50414	methylobaci
81	32	52.5	259	1	TRMB_AZOSE	Q8g9c6	azoarcus sp
82	32	52.5	345	1	GBA4_DICDI	P34042	dictyosteli
83	32	52.5	355	1	GBQ_GEOCY	Q9xzv4	geodia cydo
84	32	52.5	372	1	WNT1_CAEEL	P34888	caenorhabdi
85	32	52.5	415	1	YBDG_ECOLI	P39455	escherichia
86	32	52.5	435	1	ERF1_POLMI	Q9gr88	polyandroca
87	32	52.5	437	1	ERF1_HUMAN	P46055	homo sapien
88	32	52.5	437	1	ERF1_XENLA	P35615	xenopus lae
89	32	52.5	492	1	CPAC_MOUSE	P56593	mus musculu
90	32	52.5	531	1	MGLC_TREPA	Q57321	treponema p
91	32	52.5	554	1	Y478_RICPR	Q9zd66	rickettsia
92	32	52.5	742	1	ZW10_ARATH	O48626	arabidopsis
93	32	52.5	776	1	KLP1_CHLRE	P46870	chlamydomon
94	32	52.5	791	1	RE10_SCHPO	Q09823	schizosacch
95	32	52.5	863	1	SIP1_YEAST	P32578	saccharomyc
96	32	52.5	894	1	OPLA_HUMAN	O14841	homo sapien
97	32	52.5	1312	1	DPOL_PYRSD	Q51334	pyrococcus
98	32	52.5	1546	1	ACC9_MOUSE	P70170	mus musculu
99	32	52.5	1671	1	DPOL_PYRKO	P77933	pyrococcus
100	32	52.5	1699	1	DPOL_THEG8	Q9hh84	thermococcu
101	31	50.8	155	1	MLC1_DROPS	Q24621	drosophila
102	31	50.8	180	1	KDOP_HAEIN	P45314	haemophilus
103	31	50.8	190	1	GRPE_CHLMU	P23575	chlamydia m
104	31	50.8	195	1	MSA2_RHILO	Q98dv6	rhizobium l
105	31	50.8	214	1	KGUA_PSESM	Q88be2	pseudomonas
106	31	50.8	237	1	VD03_VARV	P33068	variola vir
107	31	50.8	240	1	HLYA_TREHY	Q06803	treponema h
108	31	50.8	249	1	YG61_FUSNN	Q8rie0	fusobacteri
109	31	50.8	257	1	Y365_WIGBR	Q8d2i9	wiggleswort
110	31	50.8	265	1	FAD1_SCHPO	O74841	schizosacch
111	31	50.8	266	1	ETC1_STAAU	P01553	staphylococ
112	31	50.8	266	1	ETXB_STAAU	P01552	staphylococ
113	31	50.8	295	1	LIPA_AERPE	Q9y9e3	aeropyrum p
114	31	50.8	298	1	RT03_ACACA	P46754	acanthamoeb
115	31	50.8	312	1	TRUB_BUCAP	Q8k9h3	buchnera ap
116	31	50.8	326	1	AOX3_SOYBN	O03376	glycine max
117	31	50.8	327	1	YRB2_YEAST	P40517	saccharomyc
118	31	50.8	332	1	OTCC_BACCR	Q81ii0	bacillus ce
119	31	50.8	353	1	GB0_XENLA	P10825	xenopus lae
120	31	50.8	355	1	GBI5_DROME	P20353	drosophila
121	31	50.8	356	1	Y359_AQUAE	O66685	aquifex aeo
122	31	50.8	368	1	P37_MYCGE	Q49410	mycoplasma
123	31	50.8	397	1	CD61_SULSO	Q980n4	sulfolobus
124	31	50.8	421	1	AK_MYCTU	P97048	mycobacteri
125	31	50.8	434	1	ERFB_ARATH	Q9lpv8	arabidopsis
126	31	50.8	435	1	ERFC_ARATH	P35614	arabidopsis
127	31	50.8	438	1	ILR3_ARATH	P54969	arabidopsis
128	31	50.8	439	1	ILL2_ARATH	P54970	arabidopsis
129	31	50.8	450	1	LIPP_PIG	P00591	sus scrofa
130	31	50.8	465	1	LIPP_CAVPO	P50903	cavia porce
131	31	50.8	465	1	LIPP_RAT	P27657	rattus norv

132	31	50.8	512	1	UGTB_CAEEL	Q22180	caenorhabdi
133	31	50.8	522	1	C5P3_ARATH	Q9fg23	arabidopsis
134	31	50.8	526	1	MVIN_TREPA	O83529	treponema p
135	31	50.8	529	1	KPYK_CHICK	P00548	gallus gall
136	31	50.8	543	1	SUW2_HUMAN	Q86yh2	homo sapien
137	31	50.8	562	1	ATKA_YERPE	Q8zd96	yersinia pe
138	31	50.8	578	1	PRIM_BUCBP	Q89b09	buchnera ap
139	31	50.8	589	1	SYFB_DROME	Q9vca5	drosophila
140	31	50.8	608	1	KU70_MOUSE	P23475	mus musculu
141	31	50.8	614	1	SPAT_BACSU	P33116	bacillus su
142	31	50.8	619	1	CYG1_BOVIN	P16068	bos taurus
143	31	50.8	619	1	CYG1_HUMAN	Q02153	homo sapien
144	31	50.8	619	1	CYG1_RAT	P20595	rattus norv
145	31	50.8	700	1	V018_FOWPV	Q9j5i3	fowlpox vir
146	31	50.8	703	1	LAGD_LACLA	P59852	lactococcus
147	31	50.8	715	1	ZP2_CANFA	P47983	canis famil
148	31	50.8	716	1	DVL3_HUMAN	Q92997	homo sapien
149	31	50.8	716	1	DVL3_MOUSE	Q61062	mus musculu
150	31	50.8	732	1	YM8K_YEAST	Q03254	saccharomyc

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.

AC Q28178; Q28179;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Thrombospondin 1 precursor.

GN THBS1 OR TSP1 OR TSP-1.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Holstein; TISSUE=Tooth;

RX MEDLINE=98173773; PubMed=9507054;

RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,

RA Inoue H.;

RT "cDNA cloning of bovine thrombospondin 1 and its expression in

RT odontoblasts and predentin.";

RL Biochim. Biophys. Acta 1382:17-22(1998).

RN [2]

RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.

RC TISSUE=Aortic endothelium;

RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;

RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.

CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and

CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,

CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-

CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis

CC and/or maintenance of dentin and dental pulp.

```

CC  -!- SUBUNIT: Homotrimer; disulfide-linked.
CC  -!- TISSUE SPECIFICITY: Odontoblasts.
CC  -!- SIMILARITY: Belongs to the thrombospondin family.
CC  -!- SIMILARITY: Contains 1 VWFC domain.
CC  -!- SIMILARITY: Contains 3 EGF-like domains.
CC  -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC  -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC  -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AB005287; BAA21115.1; -.
DR  EMBL; X87618; CAA60950.1; -.
DR  EMBL; X87619; CAA60951.1; -.
DR  PIR; S55501; S55501.
DR  GlycoSuiteDB; Q28178; -.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR006209; EGF_like.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000884; TSP1.
DR  InterPro; IPR008085; TSP_1.
DR  InterPro; IPR003367; tsp_3.
DR  InterPro; IPR008859; TSPC.
DR  InterPro; IPR003129; TSPN.
DR  InterPro; IPR001007; VWF_C.
DR  Pfam; PF00008; EGF; 2.
DR  Pfam; PF00090; tsp_1; 3.
DR  Pfam; PF02412; tsp_3; 13.
DR  Pfam; PF05735; TSPC; 1.
DR  Pfam; PF02210; TSPN; 1.
DR  Pfam; PF00093; vwc; 1.
DR  PRINTS; PR01705; TSP1REPEAT.
DR  SMART; SM00181; EGF; 3.
DR  SMART; SM00209; TSP1; 3.
DR  SMART; SM00210; TSPN; 1.
DR  SMART; SM00214; VWC; 1.
DR  PROSITE; PS00022; EGF_1; FALSE_NEG.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS50026; EGF_3; 2.
DR  PROSITE; PS50092; TSP1; 3.
DR  PROSITE; PS01208; VWFC_1; 1.
DR  PROSITE; PS50184; VWFC_2; 1.
KW  Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW  EGF-like domain; Signal.
FT  SIGNAL          1      18      BY SIMILARITY.
FT  CHAIN           19    1170    THROMBOSPONDIN 1.
FT  DOMAIN          19    232    HEPARIN-BINDING (POTENTIAL).
FT  DOMAIN          24    221    TSP N-TERMINAL.
FT  DOMAIN          316   373    VWFC.
FT  DOMAIN          379   429    TSP TYPE-1 1.
FT  DOMAIN          435   490    TSP TYPE-1 2.

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:33:49 ; Search time 31.2208 Seconds  
 (without alignments)  
 121.272 Million cell updates/sec

Title: US-10-030-735-23  
 Perfect score: 61  
 Sequence: 1 FQGV LQNVKFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertibrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description
-----						



1	58	95.1	229	6	Q28194	Q28194 bos taurus
2	58	95.1	496	13	Q7SY84	Q7sy84 xenopus lae
3	58	95.1	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	58	95.1	1171	11	Q80YQ1	Q80yq1 mus musculu
5	45	73.8	727	17	Q8PRY3	Q8pry3 methanosarc
6	45	73.8	733	17	Q8TLX6	Q8tlx6 methanosarc
7	42	68.9	1172	11	Q8CG21	Q8cg21 mus musculu
8	42	68.9	1172	11	Q7TMT3	Q7tmt3 mus musculu
9	41	67.2	146	17	Q976Q1	Q976q1 sulfolobus
10	40	65.6	223	10	Q9LMB5	Q9lmb5 arabidopsis
11	40	65.6	595	5	Q86NR6	Q86nr6 drosophila
12	40	65.6	721	5	Q9VTH0	Q9vth0 drosophila
13	39	63.9	100	16	Q83E24	Q83e24 coxiella bu
14	39	63.9	258	6	Q9BEZ5	Q9bez5 choloepus d
15	39	63.9	258	6	Q9BEZ6	Q9bez6 choloepus h
16	39	63.9	258	6	Q9BEZ4	Q9bez4 euphractus
17	39	63.9	1711	9	Q8LTK2	Q8ltk2 lactococcus
18	39	63.9	1713	9	Q94MA1	Q94ma1 lactococcus
19	38	62.3	232	5	Q8ISI2	Q8isi2 oxytricha l
20	38	62.3	460	10	Q94DF6	Q94df6 oryza sativ
21	38	62.3	740	16	Q97Q99	Q97q99 streptococc
22	37	60.7	101	1	Q8NKU9	Q8nku9 acidianus a
23	37	60.7	337	16	Q8CZC1	Q8czc1 streptococc
24	37	60.7	378	12	Q91FM2	Q91fm2 chilo iride
25	37	60.7	387	16	Q82XE5	Q82xe5 nitrosomona
26	37	60.7	431	5	O17199	O17199 caenorhabdi
27	37	60.7	546	16	Q81FG1	Q81fg1 bacillus ce
28	37	60.7	546	16	Q81SI7	Q81si7 bacillus an
29	37	60.7	647	16	Q88QB0	Q88qb0 pseudomonas
30	37	60.7	659	16	Q8RCI0	Q8rci0 thermoanaer
31	37	60.7	689	16	Q8R808	Q8r808 thermoanaer
32	37	60.7	773	13	Q8UV20	Q8uv20 sphoeroides
33	37	60.7	834	3	O43048	O43048 schizosacch
34	37	60.7	855	13	Q802A5	Q802a5 fugu rubrip
35	36	59.0	157	11	Q8BSY7	Q8bsy7 mus musculu
36	36	59.0	163	16	Q7VKL4	Q7vkl4 haemophilus
37	36	59.0	298	10	Q8L9J3	Q8l9j3 arabidopsis
38	36	59.0	343	11	Q9D7B3	Q9d7b3 mus musculu
39	36	59.0	354	5	O17701	O17701 caenorhabdi
40	36	59.0	354	13	Q9DG27	Q9dg27 gallus gall
41	36	59.0	358	13	Q8QGY3	Q8qgy3 fugu rubrip
42	36	59.0	372	11	Q80X34	Q80x34 mus musculu
43	36	59.0	750	11	Q8K251	Q8k251 mus musculu
44	36	59.0	893	16	Q9S258	Q9s258 streptomyce
45	36	59.0	939	16	Q837R8	Q837r8 enterococcu
46	36	59.0	1009	5	Q868S0	Q868s0 anopheles g
47	36	59.0	1034	5	Q7Z2B9	Q7z2b9 trypanosoma
48	36	59.0	1142	3	Q9UT41	Q9ut41 schizosacch
49	36	59.0	1290	5	Q8IER9	Q8ier9 plasmodium
50	36	59.0	3712	5	Q9VRW0	Q9vrw0 drosophila
51	35.5	58.2	1554	5	Q7YY16	Q7yy16 cryptospori
52	35	57.4	82	10	Q9ZUI2	Q9zui2 arabidopsis
53	35	57.4	106	17	Q97Z50	Q97z50 sulfolobus
54	35	57.4	145	5	Q93511	Q93511 caenorhabdi
55	35	57.4	172	16	Q81FM9	Q81fm9 bacillus ce
56	35	57.4	192	13	Q8JFS2	Q8jfs2 brachydanio
57	35	57.4	207	8	Q9MJE0	Q9mje0 pauesia pin

58	35	57.4	207	8	Q9MJD9	Q9mjd9 pauesia sil
59	35	57.4	207	8	Q9MJD8	Q9mjd8 pauesia sil
60	35	57.4	210	11	Q9D9C0	Q9d9c0 mus musculu
61	35	57.4	217	16	Q8RC20	Q8rc20 thermoanaer
62	35	57.4	223	16	Q822Z7	Q822z7 chlamydophi
63	35	57.4	242	11	Q8BM21	Q8bm21 mus musculu
64	35	57.4	298	10	O04314	O04314 arabidopsis
65	35	57.4	408	10	Q8RUU5	Q8ruu5 oryza sativ
66	35	57.4	435	4	Q8TCW3	Q8tcw3 homo sapien
67	35	57.4	439	12	Q85367	Q85367 variola maj
68	35	57.4	439	12	Q89091	Q89091 variola vir
69	35	57.4	439	12	Q8V538	Q8v538 monkeypox v
70	35	57.4	439	12	Q8V2X2	Q8v2x2 camelpox vi
71	35	57.4	439	12	O57177	O57177 vaccinia vi
72	35	57.4	439	12	Q89919	Q89919 variola vir
73	35	57.4	439	12	Q8JLH1	Q8jlh1 ectromelia
74	35	57.4	439	12	Q9JFE5	Q9jfe5 vaccinia vi
75	35	57.4	439	12	Q8QN08	Q8qn08 cowpox viru
76	35	57.4	439	12	Q9PXR8	Q9pxr8 variola vir
77	35	57.4	439	12	Q89121	Q89121 vaccinia vi
78	35	57.4	439	12	Q80E19	Q80e19 cowpox viru
79	35	57.4	498	16	Q99VY9	Q99vy9 staphylococ
80	35	57.4	498	16	Q8NXT1	Q8nxt1 staphylococ
81	35	57.4	556	16	Q8RD82	Q8rd82 thermoanaer
82	35	57.4	568	5	Q9NAL4	Q9nal4 caenorhabdi
83	35	57.4	581	10	Q9FHN7	Q9fhn7 arabidopsis
84	35	57.4	593	2	Q8GH66	Q8gh66 mycobacteri
85	35	57.4	610	16	Q822D4	Q822d4 chlamydophi
86	35	57.4	619	10	Q93X09	Q93x09 vigna mungo
87	35	57.4	627	16	Q87GG1	Q87gg1 vibrio para
88	35	57.4	628	16	Q92F14	Q92f14 listeria in
89	35	57.4	628	16	Q8YA90	Q8ya90 listeria mo
90	35	57.4	646	16	Q883C0	Q883c0 pseudomonas
91	35	57.4	667	13	Q7T258	Q7t258 alligator s
92	35	57.4	770	13	O93394	O93394 perca flave
93	35	57.4	865	10	Q7X6K3	Q7x6k3 oryza sativ
94	35	57.4	866	13	Q7ZT80	Q7zt80 cyprinus ca
95	35	57.4	952	5	Q964F9	Q964f9 spodoptera
96	35	57.4	962	5	Q7YZ58	Q7yz58 cryptospori
97	35	57.4	1019	10	Q9FVE7	Q9fve7 glycine max
98	35	57.4	1390	4	Q8IW34	Q8iw34 homo sapien
99	35	57.4	1684	13	Q8AYN8	Q8ayn8 cyprinus ca
100	35	57.4	1691	13	Q8AYN9	Q8ayn9 cyprinus ca
101	35	57.4	1829	16	Q8YZJ8	Q8yzj8 anabaena sp
102	35	57.4	3242	9	Q859P9	Q859p9 bacteriopha
103	35	57.4	3270	5	Q8IDB2	Q8idb2 plasmodium
104	34.5	56.6	206	2	Q8L0V1	Q8l0v1 escherichia
105	34.5	56.6	223	16	Q8FDP2	Q8fdp2 escherichia
106	34	55.7	156	3	Q12056	Q12056 saccharomyc
107	34	55.7	219	16	Q8ENF8	Q8enf8 oceanobacil
108	34	55.7	243	16	Q7VHY5	Q7vhy5 helicobacte
109	34	55.7	255	6	Q9BEY7	Q9bey7 trichechus
110	34	55.7	257	6	Q9BEY5	Q9bey5 macroscelid
111	34	55.7	258	6	Q9BEY4	Q9bey4 elephantulu
112	34	55.7	259	9	Q38410	Q38410 bacteriopha
113	34	55.7	269	16	Q8A780	Q8a780 bacteroides
114	34	55.7	273	16	Q8CN65	Q8cn65 staphylococ

115	34	55.7	308	2	Q53703	Q53703 staphylococ
116	34	55.7	308	2	Q53731	Q53731 staphylococ
117	34	55.7	308	9	Q9MBN3	Q9mbn3 staphylococ
118	34	55.7	311	2	Q93UU9	Q93uu9 staphylococ
119	34	55.7	311	16	O51622	O51622 borrelia bu
120	34	55.7	311	16	Q99T53	Q99t53 staphylococ
121	34	55.7	313	10	Q8LS83	Q8ls83 eriolobus t
122	34	55.7	314	2	O54081	O54081 staphylococ
123	34	55.7	322	13	Q90XJ7	Q90xj7 typhlonecte
124	34	55.7	323	13	Q90XJ4	Q90xj4 pachytriton
125	34	55.7	324	13	Q90XJ6	Q90xj6 latimeria m
126	34	55.7	343	3	Q9P866	Q9p866 candida alb
127	34	55.7	351	12	Q9YQW6	Q9yqw6 tomato yell
128	34	55.7	374	6	Q9TT59	Q9tt59 tadarida br
129	34	55.7	378	16	Q8F2B5	Q8f2b5 leptospira
130	34	55.7	425	5	Q9TXT7	Q9txt7 caenorhabdi
131	34	55.7	429	5	Q7YYV1	Q7yyv1 cryptospori
132	34	55.7	435	16	Q8R6A9	Q8r6a9 fusobacteri
133	34	55.7	459	2	O84925	O84925 streptococc
134	34	55.7	459	16	Q97PX1	Q97px1 streptococc
135	34	55.7	459	16	Q8DP70	Q8dp70 streptococc
136	34	55.7	481	5	Q18255	Q18255 caenorhabdi
137	34	55.7	504	5	Q9VAA4	Q9vaa4 drosophila
138	34	55.7	505	17	O28945	O28945 archaeoglob
139	34	55.7	508	8	Q85X81	Q85x81 zizania lat
140	34	55.7	515	3	Q11120	Q11120 schizosacch
141	34	55.7	568	3	P78947	P78947 schizosacch
142	34	55.7	605	3	Q9HGP1	Q9hgp1 schizosacch
143	34	55.7	638	13	Q7T1I5	Q7t1i5 ceryle torq
144	34	55.7	750	11	Q8BVV4	Q8bvv4 mus musculu
145	34	55.7	855	4	Q96CC6	Q96cc6 homo sapien
146	34	55.7	855	4	Q9H6E1	Q9h6e1 homo sapien
147	34	55.7	856	11	Q8VIK0	Q8vik0 mus musculu
148	34	55.7	862	4	Q96S34	Q96s34 homo sapien
149	34	55.7	903	12	Q64928	Q64928 african hor
150	34	55.7	958	2	Q845K9	Q845k9 bacillus me

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.

AC Q28194;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Thrombospondin-1 (Fragment).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lec\_gl.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 95.1%; Score 58; DB 6; Length 229;  
 Best Local Similarity 91.7%; Pred. No. 0.005;  
 Matches 11; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 FQGVQLQNVKFVF 12  
 |||||:|  
 Db 190 FQGVQLQNVRFVF 201

## RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:40:26 ; Search time 12.8831 Seconds  
 (without alignments)  
 48.087 Million cell updates/sec

Title: US-10-030-735-24  
 Perfect score: 62  
 Sequence: 1 FQGVLLNNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	56	90.3	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	38	61.3	1172	1	US-08-313-288B-19	Sequence 19, Appl
3	37	59.7	740	4	US-09-134-000C-6441	Sequence 6441, Ap
4	36	58.1	175	4	US-09-252-991A-18826	Sequence 18826, A
5	36	58.1	1288	3	US-08-762-428A-6	Sequence 6, Appli
6	35	56.5	30	2	US-08-404-531B-37	Sequence 37, Appl
7	35	56.5	30	3	US-08-476-900A-37	Sequence 37, Appl
8	35	56.5	30	3	US-08-488-546A-37	Sequence 37, Appl
9	35	56.5	65	4	US-09-540-236-3264	Sequence 3264, Ap
10	35	56.5	181	3	US-09-305-640-4	Sequence 4, Appli
11	35	56.5	640	2	US-08-671-978A-10	Sequence 10, Appl

12	35	56.5	775	3	US-09-305-640-2	Sequence 2, Appli
13	35	56.5	994	4	US-09-543-681A-7288	Sequence 7288, Ap
14	35	56.5	1498	2	US-08-404-531B-28	Sequence 28, Appl
15	35	56.5	1498	2	US-08-404-531B-29	Sequence 29, Appl
16	35	56.5	1498	3	US-08-476-900A-28	Sequence 28, Appl
17	35	56.5	1498	3	US-08-476-900A-29	Sequence 29, Appl
18	35	56.5	1498	3	US-08-488-546A-28	Sequence 28, Appl
19	35	56.5	1498	3	US-08-488-546A-29	Sequence 29, Appl
20	35	56.5	1580	3	US-08-726-320-1	Sequence 1, Appli
21	35	56.5	1580	3	US-09-208-716-1	Sequence 1, Appli
22	35	56.5	1581	2	US-08-404-531B-6	Sequence 6, Appli
23	35	56.5	1581	3	US-08-476-900A-6	Sequence 6, Appli
24	35	56.5	1581	3	US-08-488-546A-6	Sequence 6, Appli
25	35	56.5	1581	3	US-08-726-320-3	Sequence 3, Appli
26	35	56.5	1581	3	US-08-726-320-4	Sequence 4, Appli
27	35	56.5	1581	3	US-09-208-716-3	Sequence 3, Appli
28	35	56.5	1581	3	US-09-208-716-4	Sequence 4, Appli
29	35	56.5	1582	2	US-08-404-531B-9	Sequence 9, Appli
30	35	56.5	1582	3	US-08-476-900A-9	Sequence 9, Appli
31	35	56.5	1582	3	US-08-488-546A-9	Sequence 9, Appli
32	35	56.5	1582	3	US-08-726-320-5	Sequence 5, Appli
33	35	56.5	1582	3	US-09-208-716-5	Sequence 5, Appli
34	35	56.5	2987	2	US-08-970-269A-29	Sequence 29, Appl
35	35	56.5	2987	3	US-09-407-562-29	Sequence 29, Appl
36	35	56.5	3959	2	US-08-970-269A-30	Sequence 30, Appl
37	35	56.5	3959	3	US-09-407-562-30	Sequence 30, Appl
38	34	54.8	25	3	US-09-199-290-1	Sequence 1, Appli
39	34	54.8	25	4	US-09-821-616-1	Sequence 1, Appli
40	34	54.8	75	4	US-09-134-000C-4811	Sequence 4811, Ap
41	34	54.8	390	3	US-08-650-766-7	Sequence 7, Appli
42	34	54.8	390	3	US-08-922-635-6	Sequence 6, Appli
43	34	54.8	390	4	US-09-389-487-7	Sequence 7, Appli
44	34	54.8	414	4	US-09-107-532A-6114	Sequence 6114, Ap
45	34	54.8	539	4	US-09-800-170-16	Sequence 16, Appl
46	34	54.8	559	4	US-09-364-206-47	Sequence 47, Appl
47	34	54.8	591	3	US-09-199-290-7	Sequence 7, Appli
48	34	54.8	591	4	US-09-821-616-7	Sequence 7, Appli
49	34	54.8	618	3	US-09-199-290-34	Sequence 34, Appl
50	34	54.8	618	4	US-09-821-616-34	Sequence 34, Appl
51	34	54.8	651	3	US-08-650-766-6	Sequence 6, Appli
52	34	54.8	651	3	US-08-922-635-5	Sequence 5, Appli
53	34	54.8	651	4	US-09-389-487-6	Sequence 6, Appli
54	34	54.8	1019	1	US-08-271-364A-7	Sequence 7, Appli
55	34	54.8	1019	2	US-08-222-715B-26	Sequence 26, Appl
56	34	54.8	1070	3	US-08-922-635-22	Sequence 22, Appl
57	34	54.8	1504	4	US-09-364-206-2	Sequence 2, Appli
58	33	53.2	65	4	US-09-107-532A-4450	Sequence 4450, Ap
59	33	53.2	104	4	US-09-621-976-6750	Sequence 6750, Ap
60	33	53.2	151	4	US-09-134-000C-3770	Sequence 3770, Ap
61	33	53.2	169	4	US-09-134-000C-3511	Sequence 3511, Ap
62	33	53.2	226	3	US-09-176-657-3	Sequence 3, Appli
63	33	53.2	226	4	US-09-421-299-3	Sequence 3, Appli
64	33	53.2	275	4	US-09-083-268-18	Sequence 18, Appl
65	33	53.2	325	4	US-09-107-532A-4443	Sequence 4443, Ap
66	33	53.2	334	4	US-09-134-000C-5726	Sequence 5726, Ap
67	33	53.2	894	4	US-09-134-000C-5731	Sequence 5731, Ap
68	33	53.2	902	4	US-09-107-532A-3837	Sequence 3837, Ap

69	33	53.2	1272	4	US-09-543-681A-5732	Sequence 5732, Ap
70	33	53.2	2475	3	US-09-413-814-48	Sequence 48, Appl
71	32.5	52.4	101	3	US-09-100-802-9	Sequence 9, Appli
72	32.5	52.4	102	4	US-09-149-476-366	Sequence 366, App
73	32	51.6	22	1	US-08-019-073-11	Sequence 11, Appl
74	32	51.6	22	5	PCT-US94-01768-11	Sequence 11, Appl
75	32	51.6	153	2	US-08-387-942C-51	Sequence 51, Appl
76	32	51.6	175	3	US-09-230-637-24	Sequence 24, Appl
77	32	51.6	187	4	US-09-134-001C-3251	Sequence 3251, Ap
78	32	51.6	275	4	US-09-134-001C-3732	Sequence 3732, Ap
79	32	51.6	357	4	US-09-120-051D-9	Sequence 9, Appli
80	32	51.6	374	4	US-09-442-349A-7	Sequence 7, Appli
81	32	51.6	374	4	US-09-442-349A-17	Sequence 17, Appl
82	32	51.6	374	4	US-09-442-349A-18	Sequence 18, Appl
83	32	51.6	374	4	US-09-442-349A-19	Sequence 19, Appl
84	32	51.6	374	4	US-09-442-349A-22	Sequence 22, Appl
85	32	51.6	374	4	US-09-442-349A-32	Sequence 32, Appl
86	32	51.6	374	4	US-09-442-349A-33	Sequence 33, Appl
87	32	51.6	374	4	US-09-442-349A-34	Sequence 34, Appl
88	32	51.6	432	4	US-09-489-039A-7909	Sequence 7909, Ap
89	32	51.6	473	4	US-09-252-991A-29434	Sequence 29434, A
90	32	51.6	475	4	US-09-107-532A-3866	Sequence 3866, Ap
91	32	51.6	497	4	US-09-634-238-286	Sequence 286, App
92	32	51.6	533	1	US-08-484-493-13	Sequence 13, Appl
93	32	51.6	533	1	US-08-484-494-13	Sequence 13, Appl
94	32	51.6	533	2	US-08-345-212-13	Sequence 13, Appl
95	32	51.6	533	3	US-09-249-003-13	Sequence 13, Appl
96	32	51.6	533	4	US-09-685-844-13	Sequence 13, Appl
97	32	51.6	538	4	US-09-489-039A-8363	Sequence 8363, Ap
98	32	51.6	538	4	US-09-134-000C-4331	Sequence 4331, Ap
99	32	51.6	665	2	US-08-846-762-17	Sequence 17, Appl
100	32	51.6	665	2	US-08-846-762-94	Sequence 94, Appl
101	32	51.6	671	4	US-09-252-991A-28798	Sequence 28798, A
102	32	51.6	872	2	US-08-387-942C-5	Sequence 5, Appli
103	32	51.6	1022	1	US-08-271-364A-8	Sequence 8, Appli
104	32	51.6	1022	2	US-08-222-715B-27	Sequence 27, Appl
105	32	51.6	1112	3	US-09-353-585-2	Sequence 2, Appli
106	32	51.6	1112	3	US-09-353-585-3	Sequence 3, Appli
107	32	51.6	2008	4	US-09-091-501B-8	Sequence 8, Appli
108	32	51.6	3433	4	US-09-091-501B-10	Sequence 10, Appl
109	31	50.0	10	4	US-09-270-542-30	Sequence 30, Appl
110	31	50.0	24	2	US-08-404-531B-34	Sequence 34, Appl
111	31	50.0	24	3	US-08-476-900A-34	Sequence 34, Appl
112	31	50.0	24	3	US-08-488-546A-34	Sequence 34, Appl
113	31	50.0	25	2	US-08-404-531B-38	Sequence 38, Appl
114	31	50.0	25	2	US-08-404-531B-39	Sequence 39, Appl
115	31	50.0	25	3	US-08-476-900A-38	Sequence 38, Appl
116	31	50.0	25	3	US-08-476-900A-39	Sequence 39, Appl
117	31	50.0	25	3	US-08-488-546A-38	Sequence 38, Appl
118	31	50.0	25	3	US-08-488-546A-39	Sequence 39, Appl
119	31	50.0	44	4	US-09-161-939A-16	Sequence 16, Appl
120	31	50.0	104	4	US-09-328-352-7867	Sequence 7867, Ap
121	31	50.0	109	1	US-08-411-635-2	Sequence 2, Appli
122	31	50.0	109	4	US-08-845-011-2	Sequence 2, Appli
123	31	50.0	109	5	PCT-US94-10644-2	Sequence 2, Appli
124	31	50.0	112	4	US-09-621-976-4097	Sequence 4097, Ap
125	31	50.0	136	4	US-08-826-134-15	Sequence 15, Appl

126	31	50.0	136	4	US-08-826-134-26	Sequence 26, Appl
127	31	50.0	206	4	US-09-252-991A-20515	Sequence 20515, A
128	31	50.0	220	4	US-09-489-039A-9389	Sequence 9389, Ap
129	31	50.0	228	1	US-08-167-035-47	Sequence 47, Appl
130	31	50.0	228	1	US-08-208-887A-47	Sequence 47, Appl
131	31	50.0	228	2	US-08-539-005-47	Sequence 47, Appl
132	31	50.0	228	2	US-08-815-176-5	Sequence 5, Appli
133	31	50.0	228	4	US-09-280-598-44	Sequence 44, Appl
134	31	50.0	228	4	US-09-197-344-5	Sequence 5, Appli
135	31	50.0	278	2	US-08-701-191A-39	Sequence 39, Appl
136	31	50.0	278	4	US-09-664-526-39	Sequence 39, Appl
137	31	50.0	315	4	US-09-328-352-5948	Sequence 5948, Ap
138	31	50.0	326	4	US-09-083-268-17	Sequence 17, Appl
139	31	50.0	341	4	US-09-489-039A-12069	Sequence 12069, A
140	31	50.0	358	4	US-09-107-532A-4677	Sequence 4677, Ap
141	31	50.0	362	2	US-09-080-897-6	Sequence 6, Appli
142	31	50.0	362	3	US-09-323-735-6	Sequence 6, Appli
143	31	50.0	418	4	US-09-648-281-12	Sequence 12, Appl
144	31	50.0	418	4	US-09-083-268-5	Sequence 5, Appli
145	31	50.0	419	3	US-09-252-292C-25	Sequence 25, Appl
146	31	50.0	419	3	US-09-554-225-1	Sequence 1, Appli
147	31	50.0	427	4	US-09-550-645-2	Sequence 2, Appli
148	31	50.0	446	4	US-09-199-637A-267	Sequence 267, App
149	31	50.0	448	4	US-09-107-532A-6632	Sequence 6632, Ap
150	31	50.0	471	4	US-09-711-164-444	Sequence 444, App

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and AviHu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.



; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1170 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 US-08-313-288B-20

Query Match 90.3%; Score 56; DB 1; Length 1170;  
 Best Local Similarity 91.7%; Pred. No. 0.038;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLLNNVRFVF 12  
 ||||| |||||  
 Db 208 FQGVLLQNVRFVF 219

RESULT 2

US-08-313-288B-19  
 ; Sequence 19, Application US/08313288B  
 ; Patent No. 5750502  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Jessell, Thomas M. and AviHu Klar  
 ; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A  
 ; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Cooper & Dunham LLP  
 ; STREET: 1185 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/313,288B  
 ; FILING DATE: January 5, 1995  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: White, John P.  
 ; REGISTRATION NUMBER: 28,678  
 ; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 278-0400  
 ; TELEFAX: (212) 391-0526  
 ; TELEX:

OM protein - protein search, using sw model

Run on: April 7, 2004, 20:55:10 ; Search time 33.1429 Seconds  
 (without alignments)  
 95.091 Million cell updates/sec

Title: US-10-030-735-24  
 Perfect score: 62  
 Sequence: 1 FQGVLLNNRVFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	56	90.3	240	12	US-10-419-462-40	Sequence 40, Appl
2	56	90.3	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	56	90.3	831	12	US-09-939-853A-97	Sequence 97, Appl
4	56	90.3	831	12	US-09-939-853A-98	Sequence 98, Appl
5	56	90.3	1152	9	US-09-919-603-1	Sequence 1, Appli
6	56	90.3	1170	12	US-10-211-462-38	Sequence 38, Appl
7	56	90.3	1170	12	US-10-231-956A-482	Sequence 482, App
8	56	90.3	1170	12	US-10-419-462-38	Sequence 38, Appl
9	56	90.3	1170	14	US-10-020-141-12	Sequence 12, Appl
10	56	90.3	1170	14	US-10-017-721-2	Sequence 2, Appli
11	56	90.3	1170	14	US-10-021-660-114	Sequence 114, App
12	56	90.3	1170	14	US-10-008-093-2	Sequence 2, Appli
13	56	90.3	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	40	64.5	15	15	US-10-285-394-153	Sequence 153, App
15	40	64.5	597	12	US-10-282-122A-76883	Sequence 76883, A
16	39	62.9	432	12	US-10-424-599-228725	Sequence 228725,
17	38	61.3	16	9	US-09-822-682-6	Sequence 6, Appli
18	38	61.3	151	9	US-09-925-299-1251	Sequence 1251, Ap
19	38	61.3	151	10	US-09-925-299-1251	Sequence 1251, Ap
20	38	61.3	304	15	US-10-297-022-5	Sequence 5, Appli
21	38	61.3	1168	9	US-09-919-603-2	Sequence 2, Appli
22	38	61.3	1172	9	US-09-919-770-4	Sequence 4, Appli
23	38	61.3	1172	9	US-09-822-682-2	Sequence 2, Appli
24	38	61.3	1172	10	US-09-866-050A-658	Sequence 658, App
25	38	61.3	1172	14	US-10-060-036-171	Sequence 171, App
26	38	61.3	1172	14	US-10-020-141-14	Sequence 14, Appl
27	38	61.3	1172	14	US-10-017-724-2	Sequence 2, Appli
28	38	61.3	1172	14	US-10-301-822-203	Sequence 203, App
29	38	61.3	1172	14	US-10-008-093-4	Sequence 4, Appli
30	38	61.3	1172	14	US-10-007-781-2	Sequence 2, Appli
31	38	61.3	1172	15	US-10-372-683-16	Sequence 16, Appl
32	37	59.7	49	12	US-10-424-599-254334	Sequence 254334,
33	37	59.7	75	12	US-10-424-599-283340	Sequence 283340,
34	37	59.7	883	12	US-10-282-122A-61001	Sequence 61001, A
35	37	59.7	920	12	US-10-282-122A-53956	Sequence 53956, A
36	37	59.7	943	12	US-10-282-122A-73717	Sequence 73717, A
37	37	59.7	954	9	US-09-815-242-13192	Sequence 13192, A
38	37	59.7	954	9	US-09-815-242-13583	Sequence 13583, A
39	36	58.1	48	12	US-10-424-599-197249	Sequence 197249,
40	36	58.1	79	12	US-10-424-599-238347	Sequence 238347,
41	36	58.1	140	12	US-10-424-599-166847	Sequence 166847,
42	36	58.1	168	9	US-09-764-864-1223	Sequence 1223, Ap
43	36	58.1	212	12	US-10-425-114-38385	Sequence 38385, A
44	36	58.1	246	15	US-10-369-493-20585	Sequence 20585, A
45	36	58.1	710	12	US-10-282-122A-75098	Sequence 75098, A
46	35	56.5	52	12	US-10-424-599-268809	Sequence 268809,
47	35	56.5	119	11	US-09-833-245-75	Sequence 75, Appl
48	35	56.5	226	9	US-09-895-913A-362	Sequence 362, App
49	35	56.5	247	9	US-09-815-242-11342	Sequence 11342, A
50	35	56.5	247	9	US-09-815-242-11514	Sequence 11514, A
51	35	56.5	247	12	US-10-282-122A-58737	Sequence 58737, A
52	35	56.5	265	10	US-09-932-227-72	Sequence 72, Appl
53	35	56.5	265	14	US-10-183-708-72	Sequence 72, Appl
54	35	56.5	305	16	US-10-389-566-594	Sequence 594, App
55	35	56.5	449	10	US-09-798-723-4	Sequence 4, Appli
56	35	56.5	534	10	US-09-802-208B-4	Sequence 4, Appli

57	35	56.5	565	15	US-10-369-493-6976	Sequence 6976, Ap
58	35	56.5	775	10	US-09-798-723-2	Sequence 2, Appli
59	35	56.5	940	12	US-10-282-122A-52332	Sequence 52332, A
60	35	56.5	944	12	US-10-282-122A-68467	Sequence 68467, A
61	35	56.5	1305	12	US-10-282-122A-78213	Sequence 78213, A
62	35	56.5	1510	12	US-10-262-511-226	Sequence 226, App
63	35	56.5	1581	12	US-10-262-511-224	Sequence 224, App
64	35	56.5	4097	12	US-10-363-616-415	Sequence 415, App
65	35	56.5	4128	12	US-10-363-616-416	Sequence 416, App
66	34	54.8	25	10	US-09-821-616-1	Sequence 1, Appli
67	34	54.8	145	12	US-10-424-599-269584	Sequence 269584,
68	34	54.8	147	12	US-10-424-599-209120	Sequence 209120,
69	34	54.8	215	14	US-10-029-386-33750	Sequence 33750, A
70	34	54.8	226	9	US-09-925-300-1217	Sequence 1217, Ap
71	34	54.8	234	15	US-10-369-493-8765	Sequence 8765, Ap
72	34	54.8	326	9	US-09-925-300-1605	Sequence 1605, Ap
73	34	54.8	354	14	US-10-172-502-8	Sequence 8, Appli
74	34	54.8	355	13	US-10-108-605-205	Sequence 205, App
75	34	54.8	382	12	US-10-282-122A-49447	Sequence 49447, A
76	34	54.8	387	15	US-10-108-260A-3087	Sequence 3087, Ap
77	34	54.8	390	14	US-10-420-845-6	Sequence 6, Appli
78	34	54.8	406	12	US-10-282-122A-57687	Sequence 57687, A
79	34	54.8	443	16	US-10-389-566-551	Sequence 551, App
80	34	54.8	498	12	US-10-425-114-49331	Sequence 49331, A
81	34	54.8	539	15	US-10-422-536-16	Sequence 16, Appl
82	34	54.8	559	14	US-10-284-499-47	Sequence 47, Appl
83	34	54.8	584	15	US-10-369-493-19944	Sequence 19944, A
84	34	54.8	591	10	US-09-821-616-7	Sequence 7, Appli
85	34	54.8	618	10	US-09-821-616-34	Sequence 34, Appl
86	34	54.8	651	14	US-10-420-845-5	Sequence 5, Appli
87	34	54.8	943	9	US-09-815-242-10994	Sequence 10994, A
88	34	54.8	943	12	US-10-282-122A-58114	Sequence 58114, A
89	34	54.8	943	12	US-10-282-122A-67450	Sequence 67450, A
90	34	54.8	943	12	US-10-406-686A-26	Sequence 26, Appl
91	34	54.8	1070	14	US-10-420-845-22	Sequence 22, Appl
92	34	54.8	1161	12	US-10-282-122A-47195	Sequence 47195, A
93	34	54.8	1300	10	US-09-896-923-3	Sequence 3, Appli
94	34	54.8	1300	15	US-10-223-650-3	Sequence 3, Appli
95	34	54.8	1504	9	US-09-932-145-7	Sequence 7, Appli
96	34	54.8	1504	14	US-10-284-499-1	Sequence 1, Appli
97	34	54.8	1504	15	US-10-395-812-7	Sequence 7, Appli
98	34	54.8	1729	15	US-10-369-493-2510	Sequence 2510, Ap
99	34	54.8	1751	15	US-10-435-766-103	Sequence 103, App
100	33	53.2	43	12	US-10-424-599-183491	Sequence 183491,
101	33	53.2	51	12	US-10-424-599-198354	Sequence 198354,
102	33	53.2	59	11	US-09-864-408A-8466	Sequence 8466, Ap
103	33	53.2	69	12	US-10-424-599-206194	Sequence 206194,
104	33	53.2	83	12	US-10-425-114-47166	Sequence 47166, A
105	33	53.2	116	12	US-10-424-599-154870	Sequence 154870,
106	33	53.2	124	11	US-09-833-245-1905	Sequence 1905, Ap
107	33	53.2	157	12	US-10-282-122A-74285	Sequence 74285, A
108	33	53.2	177	12	US-10-424-599-240057	Sequence 240057,
109	33	53.2	186	12	US-10-282-122A-48526	Sequence 48526, A
110	33	53.2	188	9	US-09-925-300-1176	Sequence 1176, Ap
111	33	53.2	216	12	US-10-282-122A-51848	Sequence 51848, A
112	33	53.2	226	12	US-10-282-122A-76127	Sequence 76127, A
113	33	53.2	226	14	US-10-359-385-3	Sequence 3, Appli

114	33	53.2	291	12	US-10-424-599-266138	Sequence 266138,
115	33	53.2	299	12	US-10-425-114-60255	Sequence 60255, A
116	33	53.2	319	12	US-10-282-122A-72309	Sequence 72309, A
117	33	53.2	319	12	US-10-282-122A-74628	Sequence 74628, A
118	33	53.2	319	14	US-10-091-007-44	Sequence 44, Appl
119	33	53.2	321	9	US-09-815-242-10509	Sequence 10509, A
120	33	53.2	321	12	US-10-282-122A-57131	Sequence 57131, A
121	33	53.2	321	12	US-10-282-122A-57654	Sequence 57654, A
122	33	53.2	325	12	US-10-425-114-71821	Sequence 71821, A
123	33	53.2	435	12	US-10-424-599-205011	Sequence 205011,
124	33	53.2	440	12	US-10-282-122A-48298	Sequence 48298, A
125	33	53.2	461	15	US-10-369-493-8924	Sequence 8924, Ap
126	33	53.2	465	10	US-09-769-787-164	Sequence 164, App
127	33	53.2	466	12	US-10-424-599-232021	Sequence 232021,
128	33	53.2	485	15	US-10-369-493-19017	Sequence 19017, A
129	33	53.2	506	15	US-10-369-493-4743	Sequence 4743, Ap
130	33	53.2	506	15	US-10-369-493-7503	Sequence 7503, Ap
131	33	53.2	537	12	US-10-425-114-49306	Sequence 49306, A
132	33	53.2	565	12	US-10-282-122A-50809	Sequence 50809, A
133	33	53.2	772	15	US-10-369-493-23608	Sequence 23608, A
134	33	53.2	826	12	US-10-424-599-166203	Sequence 166203,
135	33	53.2	877	9	US-09-815-242-4876	Sequence 4876, Ap
136	33	53.2	880	9	US-09-815-242-10675	Sequence 10675, A
137	33	53.2	880	12	US-10-282-122A-42477	Sequence 42477, A
138	33	53.2	881	12	US-10-282-122A-57913	Sequence 57913, A
139	33	53.2	1014	12	US-10-424-599-148295	Sequence 148295,
140	33	53.2	1056	15	US-10-369-493-9664	Sequence 9664, Ap
141	33	53.2	1223	10	US-09-964-256A-2	Sequence 2, Appli
142	33	53.2	1223	14	US-10-146-733-44	Sequence 44, Appl
143	33	53.2	1273	12	US-10-282-122A-68782	Sequence 68782, A
144	33	53.2	1274	16	US-10-380-727-8	Sequence 8, Appli
145	33	53.2	3500	14	US-10-153-219-2	Sequence 2, Appli
146	33	53.2	3537	14	US-10-153-219-15	Sequence 15, Appl
147	32.5	52.4	101	10	US-09-918-036-9	Sequence 9, Appli
148	32.5	52.4	102	10	US-09-809-391-366	Sequence 366, App
149	32.5	52.4	102	10	US-09-882-171-366	Sequence 366, App
150	32.5	52.4	103	13	US-10-058-820-20	Sequence 20, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 90.3%; Score 56; DB 12; Length 240;  
Best Local Similarity 91.7%; Pred. No. 0.018;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLNNVRFVF 12  
||||| |||||  
Db 190 FQGVLLQNVRFVF 201

#### RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 90.3%; Score 56; DB 9; Length 466;  
Best Local Similarity 91.7%; Pred. No. 0.037;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FQGVLLNNVRFVF 12  
||||| |||||  
Db 261 FQGVLLQNVRFVF 272

#### RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding  
Same

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:36:09 ; Search time 10.4416 Seconds  
 (without alignments)  
 110.548 Million cell updates/sec

Title: US-10-030-735-24  
 Perfect score: 62  
 Sequence: 1 FQGVLLNNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query		DB	ID	Description
		% Match	Length			
1	56	90.3	229	2	S57957	thrombospondin 1 -
2	56	90.3	1170	1	TSHUP1	thrombospondin 1 p
3	56	90.3	1170	2	A40558	thrombospondin 1 p
4	40	64.5	597	2	B82881	hypothetical prote
5	39	62.9	311	2	F70184	ribose/galactose A
6	39	62.9	876	2	B96693	probable receptor
7	38	61.3	465	2	AC0347	probable membrane
8	38	61.3	1172	1	TSHUP2	thrombospondin 2 p
9	38	61.3	1172	2	A42587	thrombospondin 2 p
10	37	59.7	882	2	AB1631	valyl-tRNA synthet
11	37	59.7	883	2	AH1268	valyl-tRNA synthet
12	37	59.7	943	2	F95021	excinnuclease ABC,
13	37	59.7	943	2	C97893	excinnuclease ABC c

14	36	58.1	385	2	S56224	hypothetical prote
15	36	58.1	467	2	D84938	H <sup>+</sup> -transporting tw
16	36	58.1	569	2	T48484	laccase-like prote
17	36	58.1	755	2	G90095	hypothetical prote
18	36	58.1	1178	1	A39804	thrombospondin pre
19	36	58.1	1288	2	T42756	5-oxoprolinase (AT
20	36	58.1	1839	2	S77626	mannuronan C-5-epi
21	35	56.5	247	1	A64590	probable 3-oxoacyl
22	35	56.5	247	2	B71923	3-oxoacyl-[acyl-ca
23	35	56.5	271	2	C72207	hypothetical prote
24	35	56.5	316	2	S69659	hypothetical prote
25	35	56.5	440	2	S73396	hypothetical prote
26	35	56.5	565	2	T23843	hypothetical prote
27	35	56.5	603	2	S75664	sensory transducti
28	35	56.5	834	2	T39891	probable integral
29	35	56.5	1054	2	H69377	reverse gyrase (to
30	35	56.5	1071	2	T43255	tricorn proteinase
31	35	56.5	1305	2	AD0428	probable exported
32	35	56.5	1582	2	A56248	sulfonylurea recep
33	35	56.5	1829	2	AE1864	hypothetical prote
34	35	56.5	2018	2	T34274	hypothetical prote
35	35	56.5	3005	1	GNVSTV	genome polyprotein
36	35	56.5	4096	2	A57099	DNA-activated prot
37	34	54.8	111	2	S52596	probable membrane
38	34	54.8	133	2	B89969	enterotoxin Yent1
39	34	54.8	186	2	B75421	probable pilin, ty
40	34	54.8	198	2	B64363	imidazoleglycerol-
41	34	54.8	212	2	D81929	probable imidazole
42	34	54.8	262	2	A75009	probable aryl phos
43	34	54.8	304	2	T34271	hypothetical prote
44	34	54.8	306	2	B96922	transcription regu
45	34	54.8	350	2	F89883	cell surface prote
46	34	54.8	355	1	RGFFA	GTP-binding regula
47	34	54.8	397	2	E90167	hypothetical prote
48	34	54.8	397	2	AE2096	hypothetical prote
49	34	54.8	414	2	C90396	sulfide-quinone re
50	34	54.8	469	2	T33595	hypothetical prote
51	34	54.8	511	2	S77350	hypothetical prote
52	34	54.8	736	1	VPXRPC	outer layer protei
53	34	54.8	943	2	D64057	excinuclease ABC c
54	34	54.8	970	2	F64230	spore germination
55	34	54.8	993	2	T17230	hypothetical prote
56	34	54.8	1161	2	B70172	DNA polymerase III
57	34	54.8	1177	2	I64233	hypothetical prote
58	34	54.8	1312	2	S68593	DNA-directed DNA p
59	34	54.8	1670	2	S71551	DNA-directed DNA p
60	34	54.8	1729	2	T43403	1,3-beta-glucan sy
61	33	53.2	94	2	F90448	hypothetical prote
62	33	53.2	145	2	T20985	hypothetical prote
63	33	53.2	152	2	S44740	C02C2.2 protein -
64	33	53.2	162	2	B88349	protein F15D4.3 [i
65	33	53.2	196	2	F84941	amidotransferase h
66	33	53.2	210	2	A99487	hypothetical prote
67	33	53.2	216	2	D97104	uncharacterized co
68	33	53.2	219	2	D70411	hypothetical prote
69	33	53.2	226	2	AG0671	respiratory nitrat
70	33	53.2	252	2	S77108	hypothetical prote



71	33	53.2	291	2	F97135	probable membrane
72	33	53.2	300	2	D81399	malate dehydrogenase
73	33	53.2	317	2	F86901	hypothetical prote
74	33	53.2	330	2	B38265	peroxidase (EC 1.1
75	33	53.2	337	2	E97882	hypothetical prote
76	33	53.2	354	2	S27013	GTP-binding regula
77	33	53.2	363	2	AE1712	PTS system, fructo
78	33	53.2	364	2	AG1341	PTS system, fructo
79	33	53.2	376	2	AG2227	DNA repair and gen
80	33	53.2	377	2	G70143	hypothetical prote
81	33	53.2	385	2	S46532	polygalacturonase
82	33	53.2	393	2	A99361	thermostable carbo
83	33	53.2	393	2	H90291	thermostable carbo
84	33	53.2	394	2	S64933	hypothetical prote
85	33	53.2	394	2	AH0362	nucleoside permeas
86	33	53.2	401	2	E84757	hypothetical prote
87	33	53.2	403	2	T26551	hypothetical prote
88	33	53.2	456	2	H97735	hypothetical prote
89	33	53.2	482	2	S31478	alpha-amylase (EC
90	33	53.2	492	2	AG1810	glycogen (starch)
91	33	53.2	507	2	T26809	hypothetical prote
92	33	53.2	531	2	JC5172	probable methylgal
93	33	53.2	728	2	F72693	probable phosphoes
94	33	53.2	740	2	G95153	neuraminidase, pro
95	33	53.2	744	2	S45061	outer capsid spike
96	33	53.2	747	2	T40728	hypothetical prote
97	33	53.2	747	2	AE2929	two component resp
98	33	53.2	772	2	D91195	hypothetical prote
99	33	53.2	772	2	E86042	hypothetical prote
100	33	53.2	772	2	B65167	hypothetical 88.1
101	33	53.2	783	2	A98353	probable transcrip
102	33	53.2	812	1	ISZPT1	DNA topoisomerase
103	33	53.2	814	2	T50327	dna topoisomerase
104	33	53.2	926	2	AG1860	hypothetical prote
105	33	53.2	1014	2	T04721	Ca2+-transporting
106	33	53.2	1015	2	H84618	probable Ca2+-ATPa
107	33	53.2	1088	2	D82246	probable chitinase
108	33	53.2	1196	2	A29130	beta-amylase (EC 3
109	33	53.2	1247	2	T18671	hypothetical prote
110	33	53.2	1308	2	T05178	hypothetical prote
111	33	53.2	2123	2	F86348	hypothetical prote
112	33	53.2	3587	2	I40486	surfactin syntheta
113	32.5	52.4	282	2	G97101	probable nucleic a
114	32.5	52.4	1032	2	T30270	major tegumental a
115	32	51.6	84	2	T23960	hypothetical prote
116	32	51.6	109	2	S43702	major outer membra
117	32	51.6	148	2	H75096	hypothetical prote
118	32	51.6	158	2	G90361	hypothetical prote
119	32	51.6	200	2	T42547	gene 4 protein - e
120	32	51.6	210	2	AC2316	transposase alr408
121	32	51.6	212	2	T20295	hypothetical prote
122	32	51.6	242	2	T24034	hypothetical prote
123	32	51.6	250	2	D83549	hypothetical prote
124	32	51.6	272	2	T30305	dnaA protein - Lac
125	32	51.6	290	2	AD1858	indole-3-glycerol
126	32	51.6	295	2	E72462	probable lipoic ac
127	32	51.6	298	2	C87403	FdhD protein [impo

128	32	51.6	326	2	S77304	beta transducin-li
129	32	51.6	339	2	T41570	hypothetical prote
130	32	51.6	354	1	RGHUO2	GTP-binding regula
131	32	51.6	354	1	RGRT02	GTP-binding regula
132	32	51.6	354	1	RGHYO2	GTP-binding regula
133	32	51.6	354	1	RGMSO2	GTP-binding regula
134	32	51.6	354	1	RGFFO2	GTP-binding regula
135	32	51.6	354	1	RGFFO1	GTP-binding regula
136	32	51.6	354	2	S27014	GTP-binding regula
137	32	51.6	354	2	A61035	GTP-binding regula
138	32	51.6	354	2	S24362	GTP-binding regula
139	32	51.6	355	2	A48976	GTP-binding regula
140	32	51.6	363	2	S67247	hypothetical prote
141	32	51.6	371	2	T04971	hypothetical prote
142	32	51.6	371	2	C88474	protein C05D10.4 [
143	32	51.6	371	2	S68072	major outer membra
144	32	51.6	371	2	S68069	major outer membra
145	32	51.6	383	2	G96989	probable permease
146	32	51.6	385	2	S68066	major outer membra
147	32	51.6	385	2	S68067	major outer membra
148	32	51.6	385	2	S68070	major outer membra
149	32	51.6	386	2	S68062	major outer membra
150	32	51.6	417	1	D64112	glutamate-5-semial

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidias, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 90.3%; Score 56; DB 2; Length 229;  
 Best Local Similarity 91.7%; Pred. No. 0.0039;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGV LNNVRFVF 12  
 ||||| |||||  
 Db 190 FQGV LQNVRFVF 201

##### RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.

A;Reference number: A26155; MUID:87057617; PMID:2430973

A;Accession: A26155

A;Molecule type: mRNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.

A;Reference number: A34274; MUID:89291870; PMID:2544587

A;Accession: A34274

A;Molecule type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.

J. Cell Biol. 108, 729-736, 1989

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>

A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-397 <KOB>

A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276; PMID:3461443

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83, 'A', 85-374, 'RC' <DIX>

A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:22:56 ; Search time 5.61039 Seconds  
 (without alignments)  
 111.372 Million cell updates/sec

Title: US-10-030-735-24  
 Perfect score: 62  
 Sequence: 1 FQGVLLNNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

		%						
Result		Query						
No.	Score	Match	Length	DB	ID		Description	
1	56	90.3	1170	1	TSP1_BOVIN	Q28178	bos taurus	
2	56	90.3	1170	1	TSP1_HUMAN	P07996	homo sapien	
3	56	90.3	1170	1	TSP1_MOUSE	P35441	mus musculu	
4	56	90.3	1173	1	TSP1_XENLA	P35448	xenopus lae	
5	38	61.3	1172	1	TSP2_HUMAN	P35442	homo sapien	
6	38	61.3	1172	1	TSP2_MOUSE	Q03350	mus musculu	
7	37	59.7	257	1	Y365_WIGBR	Q8d2i9	wiggleswort	
8	37	59.7	943	1	UVRA_STRPN	Q97sx7	streptococc	
9	36	58.1	259	1	TRMB_AZOSE	Q8g9c6	azoarcus sp	
10	36	58.1	356	1	GBAG_CAEEL	Q9n2v6	caenorhabdi	
11	36	58.1	358	1	GBA4_CAEEL	Q9big5	caenorhabdi	
12	36	58.1	385	1	YFD0_YEAST	P43567	saccharomyc	
13	36	58.1	467	1	FLII_BUCAI	P57178	buchnera ap	
14	36	58.1	910	1	SYL_THEAC	Q9hk31	thermoplasm	
15	36	58.1	951	1	UVRA_LACPL	Q88yi7	lactobacill	
16	36	58.1	1170	1	TSP2_BOVIN	Q95116	bos taurus	
17	36	58.1	1178	1	TSP2_CHICK	P35440	gallus gall	

18	36	58.1	1288	1	OPLA_MOUSE	Q8k010	mus musculu
19	36	58.1	1288	1	OPLA_RAT	P97608	rattus norv
20	36	58.1	1839	1	ALE3_AZOVI	Q44496	azotobacter
21	35	56.5	418	1	OR13_DROME	Q9vxl0	drosophila
22	35	56.5	440	1	YA85_MYCPN	P75608	mycoplasma
23	35	56.5	775	1	ECEL_HUMAN	O95672	homo sapien
24	35	56.5	775	1	ECEL_MOUSE	Q9jmi0	mus musculu
25	35	56.5	1071	1	TRI_THEAC	P96086	thermoplasm
26	35	56.5	1580	1	ACC8_HUMAN	Q09428	homo sapien
27	35	56.5	1581	1	ACC8_CRICR	Q09427	cricetus cr
28	35	56.5	1581	1	ACC8_RAT	Q09429	rattus norv
29	35	56.5	3023	1	POLG_TVMV	P09814	t genome po
30	35	56.5	4128	1	PRKD_HUMAN	P78527	homo sapien
31	34	54.8	196	1	HIS5_METJA	Q57929	methanococc
32	34	54.8	212	1	HIS5_NEIMA	Q9jvh3	neisseria m
33	34	54.8	355	1	GBI5_DROME	P20353	drosophila
34	34	54.8	397	1	CD61_SULSO	Q980n4	sulfolobus
35	34	54.8	634	1	HWP1_CANAL	P46593	candida alb
36	34	54.8	736	1	VP4_ROTTPC	P26193	porcine rot
37	34	54.8	830	1	GCL2_MOUSE	Q9jhd2	mus musculu
38	34	54.8	837	1	GCL2_HUMAN	Q92830	homo sapien
39	34	54.8	943	1	UVRA_HAEIN	P44410	haemophilus
40	34	54.8	943	1	UVRA_PASMU	P57979	pasteurella
41	34	54.8	970	1	Y277_MYCGE	Q49409	mycoplasma
42	34	54.8	1147	1	DP3A_BORBU	O51526	borrelia bu
43	34	54.8	1177	1	Y307_MYCGE	P47549	mycoplasma
44	34	54.8	1312	1	DPOL_PYRSD	Q51334	pyrococcus
45	34	54.8	1671	1	DPOL_PYRKO	P77933	pyrococcus
46	34	54.8	1699	1	DPOL_THEG8	Q9hh84	thermococcu
47	34	54.8	1729	1	BGS1_SCHPO	Q10287	schizosacch
48	33	53.2	196	1	HIS5_BUCAI	P57204	buchnera ap
49	33	53.2	196	1	HIS5_BUCBP	P59501	buchnera ap
50	33	53.2	219	1	YC87_AQUAE	O67319	aquifex aeo
51	33	53.2	250	1	UBIE_COXBU	Q83a90	coxiella bu
52	33	53.2	330	1	PER2_ARAHY	P22196	arachis hyp
53	33	53.2	342	1	Y4WF_RHISN	P55684	rhizobium s
54	33	53.2	353	1	GBI_HELTI	P51876	helisoma tr
55	33	53.2	353	1	GBI_LYMST	P30682	lymnaea sta
56	33	53.2	376	1	RECF_ANASP	Q8yrr9	anabaena sp
57	33	53.2	393	1	CBP1_SULSO	P80092	sulfolobus
58	33	53.2	393	1	CBP2_SULSO	P58156	sulfolobus
59	33	53.2	472	1	ATPB_FERIS	O50341	fervidobact
60	33	53.2	492	1	GLG2_ANASP	Q8z0q9	anabaena sp
61	33	53.2	507	1	TX33_CAEEL	O45291	caenorhabdi
62	33	53.2	531	1	MGLC_TREPA	Q57321	treponema p
63	33	53.2	700	1	V018_FOWPV	Q9j5i3	fowlpox vir
64	33	53.2	772	1	YICI_ECOLI	P31434	escherichia
65	33	53.2	814	1	TOP1_SCHPO	P07799	schizosacch
66	33	53.2	1014	1	ACA2_ARATH	O81108	arabidopsis
67	33	53.2	1015	1	ACA7_ARATH	O64806	arabidopsis
68	33	53.2	1196	1	AMYB_PAEPO	P21543	paenibacill
69	33	53.2	1391	1	RPC1_HUMAN	O14802	homo sapien
70	33	53.2	3587	1	SRF2_BACSU	Q04747	bacillus su
71	32.5	52.4	101	1	SM33_HUMAN	Q93068	homo sapien
72	32	51.6	59	1	RL30_LEPIN	Q9xd18	leptospira
73	32	51.6	146	1	MRAZ_RHIME	Q92nl3	rhizobium m
74	32	51.6	295	1	LIPA_AERPE	Q9y9e3	aeropyrum p

75	32	51.6	326	1	YE09_SYNY3	P73594	synechocyst
76	32	51.6	353	1	GB02_CRILO	P17806	cricetulus
77	32	51.6	353	1	GB02_HUMAN	P29777	homo sapien
78	32	51.6	353	1	GB02_MOUSE	P18873	mus musculu
79	32	51.6	353	1	GB02_RAT	P30033	rattus norv
80	32	51.6	353	1	GB0_HELTI	P51877	helisoma tr
81	32	51.6	353	1	GB0_LYMST	P30683	lymnaea sta
82	32	51.6	353	1	GBI_ASTPE	P30676	asterina pe
83	32	51.6	354	1	GB01_DROME	P16378	drosophila
84	32	51.6	354	1	GB0_LOCFI	P38404	locusta mig
85	32	51.6	354	1	GBI_HOMAM	P41776	homarus ame
86	32	51.6	356	1	GB0_PATYE	O15976	patinopecte
87	32	51.6	363	1	OP2B_HAEIN	Q48220	haemophilus
88	32	51.6	367	1	OP2A_HAEIN	Q48219	haemophilus
89	32	51.6	369	1	OP27_HAEIN	Q48218	haemophilus
90	32	51.6	371	1	OP25_HAEIN	P46027	haemophilus
91	32	51.6	385	1	OP23_HAEIN	P46025	haemophilus
92	32	51.6	386	1	OP24_HAEIN	P46026	haemophilus
93	32	51.6	399	1	RMS5_PENUR	P47907	penicillium
94	32	51.6	417	1	PROA_HAEIN	P45121	haemophilus
95	32	51.6	455	1	DNAA_LACLA	Q9cjj2	lactococcus
96	32	51.6	468	1	LIP1_CANAL	O94091	candida alb
97	32	51.6	474	1	NU4M_PARTE	P15581	paramecium
98	32	51.6	494	1	ACH6_CHICK	P49581	gallus gall
99	32	51.6	504	1	COBQ_METTH	O26880	methanobact
100	32	51.6	538	1	KPYK_TRIRE	P31865	trichoderma
101	32	51.6	550	1	RT03_OENBE	P27754	oenothera b
102	32	51.6	597	1	YPC4_CAEEL	Q11181	caenorhabdi
103	32	51.6	610	1	U171_DROME	Q9vub4	drosophila
104	32	51.6	619	1	CYG1_BOVIN	P16068	bos taurus
105	32	51.6	619	1	CYG1_HUMAN	Q02153	homo sapien
106	32	51.6	619	1	CYG1_RAT	P20595	rattus norv
107	32	51.6	654	1	PSTA_MYCGE	P47651	mycoplasma
108	32	51.6	691	1	LCN3_LACLA	P37608	lactococcus
109	32	51.6	703	1	LAGD_LACLA	P59852	lactococcus
110	32	51.6	726	1	RNR_MYCPN	P75529	mycoplasma
111	32	51.6	958	1	UVRA_OCEIH	Q8enj6	oceanobacil
112	32	51.6	1020	1	ACA1_ARATH	Q37145	arabidopsis
113	32	51.6	1069	1	ACAA_ARATH	Q9szr1	arabidopsis
114	32	51.6	1220	1	ATB1_PIG	P23220	sus scrofa
115	32	51.6	1249	1	ATB1_RABIT	Q00804	oryctolagus
116	32	51.6	1258	1	ATB1_HUMAN	P20020	homo sapien
117	32	51.6	1258	1	ATB1_RAT	P11505	rattus norv
118	32	51.6	1702	1	DPOL_THELI	P30317	thermococcu
119	32	51.6	3433	1	UTRO_HUMAN	P46939	homo sapien
120	32	51.6	3712	1	LMA_DROME	Q00174	drosophila
121	32	51.6	4128	1	PRKD_MOUSE	P97313	mus musculu
122	31.5	50.8	502	1	C72A_ARATH	Q9lvd2	arabidopsis
123	31	50.0	108	1	T2AG_ONCMY	Q90yg6	oncorhynchu
124	31	50.0	109	1	T2AG_HUMAN	P52657	homo sapien
125	31	50.0	109	1	T2AG_RAT	O08950	rattus norv
126	31	50.0	111	1	T2AG_PAROL	Q9ia78	paralichthy
127	31	50.0	130	1	SAA4_MOUSE	P31532	mus musculu
128	31	50.0	155	1	YB88_AQUAE	O67248	aquifex aeo
129	31	50.0	176	1	NU6C_HORVU	O98693	hordeum vul
130	31	50.0	176	1	NU6C_MAIZE	P46621	zea mays (m
131	31	50.0	176	1	NU6C_ORYSA	P12130	oryza sativ

132	31	50.0	176	1	NU6C_WHEAT	Q95h44	triticum ae
133	31	50.0	183	1	VMO1_CHICK	P41366	gallus gall
134	31	50.0	184	1	VP50_BPAPS	Q9t1p8	bacterioph
135	31	50.0	199	1	HIS5_HAEIN	P44340	haemophilus
136	31	50.0	220	1	YQJA_ECOLI	P42614	escherichia
137	31	50.0	226	1	NARV_ECOLI	P19316	escherichia
138	31	50.0	228	1	SEM5_CAEEL	P29355	caenorhabdi
139	31	50.0	265	1	FAD1_SCHPO	O74841	schizosacch
140	31	50.0	309	1	PYRB_VIBPA	Q871f8	vibrio para
141	31	50.0	309	1	PYRB_VIBVU	Q8dcf6	vibrio vuln
142	31	50.0	326	1	AOX3_SOYBN	O03376	glycine max
143	31	50.0	333	1	LDH_CAEEL	Q27888	caenorhabdi
144	31	50.0	337	1	RBSR_PASMU	Q9cpa2	pasteurella
145	31	50.0	337	1	YA65_METJA	Q58465	methanococc
146	31	50.0	355	1	GB14_HUMAN	O95837	homo sapien
147	31	50.0	380	1	GB12_HUMAN	Q03113	homo sapien
148	31	50.0	388	1	YNJB_ECOLI	P76223	escherichia
149	31	50.0	392	1	OXDD_BACSU	O34767	bacillus su
150	31	50.0	415	1	YBDG_ECOLI	P39455	escherichia

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.  
AC Q28178; Q28179;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Thrombospondin 1 precursor.  
GN THBS1 OR TSP1 OR TSP-1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Holstein; TISSUE=Tooth;  
RX MEDLINE=98173773; PubMed=9507054;  
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,  
RA Inoue H.;  
RT "cDNA cloning of bovine thrombospondin 1 and its expression in  
RT odontoblasts and predentin.";  
RL Biochim. Biophys. Acta 1382:17-22(1998).  
RN [2]  
RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.  
RC TISSUE=Aortic endothelium;  
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
CC !- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-  
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis  
CC and/or maintenance of dentin and dental pulp.

```

CC  -!- SUBUNIT: Homotrimer; disulfide-linked.
CC  -!- TISSUE SPECIFICITY: Odontoblasts.
CC  -!- SIMILARITY: Belongs to the thrombospondin family.
CC  -!- SIMILARITY: Contains 1 VWFC domain.
CC  -!- SIMILARITY: Contains 3 EGF-like domains.
CC  -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC  -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC  -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AB005287; BAA21115.1; -.
DR  EMBL; X87618; CAA60950.1; -.
DR  EMBL; X87619; CAA60951.1; -.
DR  PIR; S55501; S55501.
DR  GlycoSuiteDB; Q28178; -.
DR  InterPro; IPR001881; EGF_Ca.
DR  InterPro; IPR006209; EGF_like.
DR  InterPro; IPR006210; IEGF.
DR  InterPro; IPR000884; TSP1.
DR  InterPro; IPR008085; TSP_1.
DR  InterPro; IPR003367; tsp_3.
DR  InterPro; IPR008859; TSPC.
DR  InterPro; IPR003129; TSPN.
DR  InterPro; IPR001007; VWF_C.
DR  Pfam; PF00008; EGF; 2.
DR  Pfam; PF00090; tsp_1; 3.
DR  Pfam; PF02412; tsp_3; 13.
DR  Pfam; PF05735; TSPC; 1.
DR  Pfam; PF02210; TSPN; 1.
DR  Pfam; PF00093; vwc; 1.
DR  PRINTS; PR01705; TSP1REPEAT.
DR  SMART; SM00181; EGF; 3.
DR  SMART; SM00209; TSP1; 3.
DR  SMART; SM00210; TSPN; 1.
DR  SMART; SM00214; VWC; 1.
DR  PROSITE; PS00022; EGF_1; FALSE_NEG.
DR  PROSITE; PS01186; EGF_2; 1.
DR  PROSITE; PS50026; EGF_3; 2.
DR  PROSITE; PS50092; TSP1; 3.
DR  PROSITE; PS01208; VWFC_1; 1.
DR  PROSITE; PS50184; VWFC_2; 1.
KW  Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW  EGF-like domain; Signal.
FT  SIGNAL      1      18      BY SIMILARITY.
FT  CHAIN       19     1170     THROMBOSPONDIN 1.
FT  DOMAIN      19     232     HEPARIN-BINDING (POTENTIAL).
FT  DOMAIN      24     221     TSP N-TERMINAL.
FT  DOMAIN     316     373     VWFC.
FT  DOMAIN     379     429     TSP TYPE-1 1.
FT  DOMAIN     435     490     TSP TYPE-1 2.

```



OM protein - protein search, using sw model

Run on: April 7, 2004, 18:33:49 ; Search time 31.2208 Seconds  
 (without alignments)  
 121.272 Million cell updates/sec

Title: US-10-030-735-24  
 Perfect score: 62  
 Sequence: 1 FQGVLLNNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp Vertebrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_cheap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%				
No.	Score	Match	Length	DB	ID	Description
-----						

1	56	90.3	229	6	Q28194	Q28194 bos taurus
2	56	90.3	496	13	Q7SY84	Q7sy84 xenopus lae
3	56	90.3	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	56	90.3	1171	11	Q80YQ1	Q80yq1 mus musculu
5	44	71.0	727	17	Q8PRY3	Q8pry3 methanosarc
6	44	71.0	733	17	Q8TLX6	Q8tlx6 methanosarc
7	42	67.7	100	16	Q83E24	Q83e24 coxiella bu
8	40	64.5	597	16	Q9PPX0	Q9ppx0 ureaplasma
9	40	64.5	659	16	Q8RCI0	Q8rci0 thermoanaer
10	40	64.5	689	16	Q8R808	Q8r808 thermoanaer
11	40	64.5	1034	5	Q7Z2B9	Q7z2b9 trypanosoma
12	39	62.9	311	16	O51622	O51622 borrelia bu
13	39	62.9	727	10	Q9FZI5	Q9fzi5 arabidopsis
14	39	62.9	876	10	Q9C622	Q9c622 arabidopsis
15	38	61.3	180	11	Q8BUC7	Q8buc7 mus musculu
16	38	61.3	223	10	Q9LMB5	Q9lmb5 arabidopsis
17	38	61.3	447	16	Q7VHM6	Q7vhm6 helicobacte
18	38	61.3	465	16	Q8ZCV8	Q8zcv8 yersinia pe
19	38	61.3	526	5	Q95UJ4	Q95uj4 aphix gossy
20	38	61.3	526	5	O76177	O76177 aphix gossy
21	38	61.3	1172	11	Q8CG21	Q8cg21 mus musculu
22	38	61.3	1172	11	Q7TMT3	Q7tmt3 mus musculu
23	37	59.7	280	10	Q94BG4	Q94bg4 froelichia
24	37	59.7	328	16	Q83DA4	Q83da4 coxiella bu
25	37	59.7	409	17	Q8TNV3	Q8tnv3 methanosarc
26	37	59.7	412	16	Q8DUJ9	Q8duj9 streptococc
27	37	59.7	423	16	Q81BG4	Q81bg4 bacillus ce
28	37	59.7	464	8	Q9BAA8	Q9baa8 melichrus p
29	37	59.7	470	8	O78404	O78404 pentachondr
30	37	59.7	474	8	Q9BAB1	Q9bab1 cyathodes g
31	37	59.7	478	8	Q9BAA1	Q9baa1 styphelia v
32	37	59.7	882	16	Q92BG2	Q92bg2 listeria in
33	37	59.7	883	16	Q8Y6X9	Q8y6x9 listeria mo
34	37	59.7	1711	9	Q8LTK2	Q8ltk2 lactococcus
35	37	59.7	1713	9	Q94MA1	Q94ma1 lactococcus
36	36	58.1	38	2	Q8KYM0	Q8kym0 bacillus an
37	36	58.1	146	9	Q8SCN4	Q8scn4 pseudomonas
38	36	58.1	146	17	Q976Q1	Q976q1 sulfolobus
39	36	58.1	166	4	Q9UDQ9	Q9udq9 homo sapien
40	36	58.1	174	2	Q84I54	Q84i54 buchnera ap
41	36	58.1	206	4	Q86WP4	Q86wp4 homo sapien
42	36	58.1	278	10	Q94BI7	Q94bi7 celosia sp.
43	36	58.1	293	10	Q84WP0	Q84wp0 arabidopsis
44	36	58.1	297	5	Q86F44	Q86f44 schistosoma
45	36	58.1	317	10	Q9FI50	Q9fi50 arabidopsis
46	36	58.1	329	3	Q8NIY1	Q8niy1 neurospora
47	36	58.1	354	2	Q9KW67	Q9kw67 staphylococ
48	36	58.1	370	17	Q8TUU8	Q8tuu8 methanopyru
49	36	58.1	373	12	Q8JME6	Q8jme6 mamestra co
50	36	58.1	391	5	Q8T3J1	Q8t3j1 drosophila
51	36	58.1	400	11	Q8K2Z1	Q8k2z1 mus musculu
52	36	58.1	533	2	Q8GDK2	Q8gdk2 photorhabdu
53	36	58.1	561	10	Q94IN2	Q94in2 hordeum vul
54	36	58.1	569	10	Q9LYQ2	Q9lyq2 arabidopsis
55	36	58.1	578	5	Q9VCG4	Q9vcg4 drosophila
56	36	58.1	586	4	Q9BQF8	Q9bqf8 homo sapien
57	36	58.1	586	4	Q8IXQ5	Q8ixq5 homo sapien

58	36	58.1	586	11	Q9CZP4	Q9czp4 mus musculu
59	36	58.1	586	11	Q8BUL5	Q8bul5 mus musculu
60	36	58.1	692	16	Q87PV7	Q87pv7 vibrio para
61	36	58.1	755	10	Q98RQ2	Q98rq2 guillardia
62	36	58.1	796	16	Q8XHL0	Q8xhl0 clostridium
63	36	58.1	951	16	Q88YI7	Q88yi7 lactobacill
64	36	58.1	1159	16	Q8D2H5	Q8d2h5 wiggleswort
65	36	58.1	9271	5	Q8IES7	Q8ies7 plasmodium
66	35	56.5	94	2	Q9ZBB3	Q9zbb3 streptococc
67	35	56.5	101	16	Q7VC17	Q7vc17 prochloroco
68	35	56.5	149	10	Q8LPX4	Q8lpx4 oryza sativ
69	35	56.5	159	10	Q7Y1P6	Q7y1p6 oryza sativ
70	35	56.5	247	16	O25286	O25286 helicobacte
71	35	56.5	247	16	Q9ZLS0	Q9zls0 helicobacte
72	35	56.5	254	16	Q7VRJ1	Q7vrj1 candidatus
73	35	56.5	258	6	Q9BEZ5	Q9bez5 choloepus d
74	35	56.5	258	6	Q9BEZ6	Q9bez6 choloepus h
75	35	56.5	258	6	Q9BEZ4	Q9bez4 euphractus
76	35	56.5	271	16	Q9X2E4	Q9x2e4 thermotoga
77	35	56.5	296	16	Q83K98	Q83k98 shigella fl
78	35	56.5	316	3	Q03419	Q03419 saccharomyc
79	35	56.5	323	2	Q937X6	Q937x6 edwardsiell
80	35	56.5	410	16	Q9CKF1	Q9ckf1 pasteurella
81	35	56.5	493	11	Q8BIU4	Q8biu4 mus musculu
82	35	56.5	534	2	Q9F4L6	Q9f4l6 escherichia
83	35	56.5	567	10	Q941X2	Q941x2 oryza sativ
84	35	56.5	573	5	Q9GQU9	Q9gqu9 caenorhabdi
85	35	56.5	603	16	P74139	P74139 synechocyst
86	35	56.5	607	2	Q9S1H6	Q9s1h6 staphylococ
87	35	56.5	733	12	Q65525	Q65525 bovine grou
88	35	56.5	834	3	O43048	O43048 schizosacch
89	35	56.5	845	4	Q86TH4	Q86th4 homo sapien
90	35	56.5	899	11	Q9Z2P2	Q9z2p2 mus musculu
91	35	56.5	942	16	Q7VLW2	Q7vlw2 haemophilus
92	35	56.5	1021	10	Q8H7X6	Q8h7x6 oryza sativ
93	35	56.5	1054	17	O29238	O29238 archaeglob
94	35	56.5	1130	16	Q8EWE7	Q8ewe7 mycoplasma
95	35	56.5	1235	16	Q7UQ00	Q7uq00 rhodopirell
96	35	56.5	1305	16	Q8ZB96	Q8zb96 yersinia pe
97	35	56.5	1829	16	Q8YZJ8	Q8yzj8 anabaena sp
98	35	56.5	2018	5	Q20487	Q20487 caenorhabdi
99	35	56.5	2232	5	Q7YYZ2	Q7yyz2 cryptospori
100	35	56.5	3023	12	Q88925	Q88925 tobacco vei
101	35	56.5	4106	6	Q8WN21	Q8wn21 equus cabal
102	35	56.5	4128	4	Q7Z611	Q7z611 homo sapien
103	34	54.8	132	2	Q9EZM6	Q9ezm6 staphylococ
104	34	54.8	132	3	O60076	O60076 schizosacch
105	34	54.8	133	16	Q99T48	Q99t48 staphylococ
106	34	54.8	168	16	Q8RIJ3	Q8rij3 fusobacteri
107	34	54.8	169	11	O70180	O70180 rattus norv
108	34	54.8	186	16	Q9RUZ7	Q9ruz7 deinococcus
109	34	54.8	190	12	Q9DHH4	Q9dhh4 yaba-like d
110	34	54.8	219	16	Q8ENF8	Q8enf8 oceanobacil
111	34	54.8	230	2	O85184	O85184 escherichia
112	34	54.8	262	17	Q9UYA1	Q9uya1 pyrococcus
113	34	54.8	306	16	Q97ML3	Q97ml3 clostridium
114	34	54.8	336	10	Q9LFP3	Q9lfp3 arabidopsis

115	34	54.8	339	5	Q9VJD6	Q9vjd6 drosophila
116	34	54.8	345	8	Q94V78	Q94v78 heloderma s
117	34	54.8	345	8	Q9TGA2	Q9tga2 heloderma s
118	34	54.8	350	16	Q99UX4	Q99ux4 staphylococ
119	34	54.8	378	8	Q9GDZ5	Q9gdz5 aeschynomen
120	34	54.8	386	16	Q7V6T2	Q7v6t2 prochloroco
121	34	54.8	387	4	Q96FK6	Q96fk6 homo sapien
122	34	54.8	395	2	O69347	O69347 vibrio sp.
123	34	54.8	395	5	Q9VK80	Q9vk80 drosophila
124	34	54.8	397	16	Q8YUL7	Q8yul7 anabaena sp
125	34	54.8	405	16	Q8F786	Q8f786 leptospira
126	34	54.8	414	17	Q97WG1	Q97wg1 sulfolobus
127	34	54.8	446	16	Q899N6	Q899n6 clostridium
128	34	54.8	468	10	Q9LP66	Q9lp66 arabidopsis
129	34	54.8	469	5	Q9TZA3	Q9tza3 caenorhabdi
130	34	54.8	487	16	Q82XZ5	Q82xz5 nitrosomona
131	34	54.8	492	16	Q89NM6	Q89nm6 bradyrhizob
132	34	54.8	493	3	O59846	O59846 aspergillus
133	34	54.8	494	13	O73621	O73621 xenopus lae
134	34	54.8	511	8	Q9GI81	Q9gi81 nissolia sc
135	34	54.8	511	8	Q9GI80	Q9gi80 nissolia hi
136	34	54.8	511	16	P73413	P73413 synechocyst
137	34	54.8	542	2	Q56294	Q56294 thiobacillu
138	34	54.8	561	10	Q851Y2	Q851y2 oryza sativ
139	34	54.8	568	5	Q9NAL4	Q9nal4 caenorhabdi
140	34	54.8	579	13	O73929	O73929 scylliorhinu
141	34	54.8	595	4	Q9UEU4	Q9ueu4 homo sapien
142	34	54.8	595	5	Q86NR6	Q86nr6 drosophila
143	34	54.8	603	16	Q88V58	Q88v58 lactobacill
144	34	54.8	603	17	Q8TL61	Q8tl61 methanosarc
145	34	54.8	618	3	Q9C1V4	Q9clv4 talaromyces
146	34	54.8	662	5	Q9VMQ8	Q9vmq8 drosophila
147	34	54.8	662	5	Q86NL4	Q86nl4 drosophila
148	34	54.8	721	5	Q9VTH0	Q9vth0 drosophila
149	34	54.8	763	3	O60013	O60013 pneumocysti
150	34	54.8	804	13	Q90ZI2	Q90zi2 gallus gall

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.

AC Q28194;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Thrombospondin-1 (Fragment).

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lectin\_g1.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 90.3%; Score 56; DB 6; Length 229;  
 Best Local Similarity 91.7%; Pred. No. 0.04;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQGVLNVRVFV 12  
 ||||| |||||  
 Db 190 FQGVLRNVRVFV 201

## RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:40:26 ; Search time 12.8831 Seconds  
 (without alignments)  
 48.087 Million cell updates/sec

Title: US-10-030-735-25  
 Perfect score: 59  
 Sequence: 1 A QGV LQNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	% Match	Query Length	DB	ID	Description
1	55	93.2	1170	1	US-08-313-288B-20	Sequence 20, Appl
2	37	62.7	446	4	US-09-199-637A-267	Sequence 267, App
3	37	62.7	1172	1	US-08-313-288B-19	Sequence 19, Appl
4	35	59.3	175	4	US-09-252-991A-18826	Sequence 18826, A
5	34	57.6	390	3	US-08-650-766-7	Sequence 7, Appli
6	34	57.6	390	3	US-08-922-635-6	Sequence 6, Appli
7	34	57.6	390	4	US-09-328-352-4891	Sequence 4891, Ap
8	34	57.6	390	4	US-09-389-487-7	Sequence 7, Appli
9	34	57.6	559	4	US-09-364-206-47	Sequence 47, Appl
10	34	57.6	651	3	US-08-650-766-6	Sequence 6, Appli
11	34	57.6	651	3	US-08-922-635-5	Sequence 5, Appli

12	34	57.6	651	4	US-09-389-487-6	Sequence 6, Appli
13	34	57.6	1070	3	US-08-922-635-22	Sequence 22, Appl
14	34	57.6	1504	4	US-09-364-206-2	Sequence 2, Appli
15	33	55.9	40	1	US-07-868-353A-3	Sequence 3, Appli
16	33	55.9	40	2	US-08-407-804-3	Sequence 3, Appli
17	33	55.9	40	3	US-09-124-807-3	Sequence 3, Appli
18	33	55.9	275	4	US-09-134-001C-3732	Sequence 3732, Ap
19	33	55.9	350	1	US-07-868-353A-14	Sequence 14, Appl
20	33	55.9	350	2	US-08-407-804-23	Sequence 23, Appl
21	33	55.9	350	3	US-09-124-807-23	Sequence 23, Appl
22	33	55.9	354	1	US-07-868-353A-12	Sequence 12, Appl
23	33	55.9	354	1	US-07-868-353A-13	Sequence 13, Appl
24	33	55.9	354	1	US-07-868-353A-15	Sequence 15, Appl
25	33	55.9	354	2	US-08-407-804-21	Sequence 21, Appl
26	33	55.9	354	2	US-08-407-804-22	Sequence 22, Appl
27	33	55.9	354	2	US-08-407-804-24	Sequence 24, Appl
28	33	55.9	354	3	US-09-124-807-21	Sequence 21, Appl
29	33	55.9	354	3	US-09-124-807-22	Sequence 22, Appl
30	33	55.9	354	3	US-09-124-807-24	Sequence 24, Appl
31	33	55.9	368	4	US-09-134-000C-4546	Sequence 4546, Ap
32	33	55.9	903	4	US-09-252-991A-28775	Sequence 28775, A
33	33	55.9	1307	1	US-08-395-246C-2	Sequence 2, Appli
34	32	54.2	24	1	US-08-019-073-15	Sequence 15, Appl
35	32	54.2	24	5	PCT-US94-01768-15	Sequence 15, Appl
36	32	54.2	170	3	US-09-370-473-2	Sequence 2, Appli
37	32	54.2	394	4	US-09-134-001C-4834	Sequence 4834, Ap
38	32	54.2	418	3	US-09-370-473-6	Sequence 6, Appli
39	32	54.2	419	3	US-09-252-292C-25	Sequence 25, Appl
40	32	54.2	476	4	US-09-489-039A-8120	Sequence 8120, Ap
41	32	54.2	740	4	US-09-134-000C-6441	Sequence 6441, Ap
42	32	54.2	747	4	US-09-724-864-36	Sequence 36, Appl
43	32	54.2	894	4	US-09-134-000C-5731	Sequence 5731, Ap
44	32	54.2	902	4	US-09-107-532A-3837	Sequence 3837, Ap
45	31	52.5	57	4	US-09-621-976-5916	Sequence 5916, Ap
46	31	52.5	171	4	US-09-328-352-8227	Sequence 8227, Ap
47	31	52.5	174	2	US-08-683-262B-53	Sequence 53, Appl
48	31	52.5	174	2	US-08-683-262B-56	Sequence 56, Appl
49	31	52.5	174	3	US-09-361-707-53	Sequence 53, Appl
50	31	52.5	174	3	US-09-361-707-56	Sequence 56, Appl
51	31	52.5	213	4	US-09-489-039A-14310	Sequence 14310, A
52	31	52.5	226	3	US-09-176-657-3	Sequence 3, Appli
53	31	52.5	226	4	US-09-421-299-3	Sequence 3, Appli
54	31	52.5	238	6	5405943-2	Patent No. 5405943
55	31	52.5	321	4	US-09-498-520A-18	Sequence 18, Appl
56	31	52.5	342	4	US-09-134-001C-4071	Sequence 4071, Ap
57	31	52.5	358	4	US-09-252-991A-27541	Sequence 27541, A
58	31	52.5	382	3	US-08-984-618-12	Sequence 12, Appl
59	31	52.5	400	4	US-09-489-039A-11916	Sequence 11916, A
60	31	52.5	406	4	US-09-543-681A-6072	Sequence 6072, Ap
61	31	52.5	415	4	US-09-489-039A-10457	Sequence 10457, A
62	31	52.5	420	4	US-09-252-991A-19389	Sequence 19389, A
63	31	52.5	436	6	5405943-4	Patent No. 5405943
64	31	52.5	565	4	US-09-134-000C-5782	Sequence 5782, Ap
65	31	52.5	864	4	US-09-134-000C-6025	Sequence 6025, Ap
66	31	52.5	911	2	US-08-928-692-59	Sequence 59, Appl
67	31	52.5	911	4	US-09-339-972-59	Sequence 59, Appl
68	31	52.5	916	2	US-08-928-692-58	Sequence 58, Appl

69	31	52.5	916	4	US-09-339-972-58	Sequence 58, Appl
70	31	52.5	1288	3	US-08-762-428A-6	Sequence 6, Appli
71	30	50.8	64	4	US-09-328-352-6751	Sequence 6751, Ap
72	30	50.8	134	4	US-09-489-039A-8594	Sequence 8594, Ap
73	30	50.8	139	4	US-09-540-236-3704	Sequence 3704, Ap
74	30	50.8	154	4	US-09-134-000C-5680	Sequence 5680, Ap
75	30	50.8	186	1	US-08-026-758-23	Sequence 23, Appl
76	30	50.8	206	4	US-09-252-991A-20515	Sequence 20515, A
77	30	50.8	259	4	US-09-252-991A-28573	Sequence 28573, A
78	30	50.8	299	4	US-09-648-004-22	Sequence 22, Appl
79	30	50.8	334	4	US-09-107-532A-6586	Sequence 6586, Ap
80	30	50.8	341	4	US-09-489-039A-12069	Sequence 12069, A
81	30	50.8	368	4	US-09-819-607-5	Sequence 5, Appli
82	30	50.8	380	4	US-09-198-452A-87	Sequence 87, Appl
83	30	50.8	385	4	US-09-134-000C-4952	Sequence 4952, Ap
84	30	50.8	409	4	US-09-489-039A-14217	Sequence 14217, A
85	30	50.8	439	1	US-08-507-431-4	Sequence 4, Appli
86	30	50.8	439	2	US-08-700-546-2	Sequence 2, Appli
87	30	50.8	439	2	US-08-902-655A-4	Sequence 4, Appli
88	30	50.8	439	3	US-09-116-622-4	Sequence 4, Appli
89	30	50.8	439	3	US-09-219-277-4	Sequence 4, Appli
90	30	50.8	439	3	US-09-599-661-4	Sequence 4, Appli
91	30	50.8	449	4	US-09-489-039A-8009	Sequence 8009, Ap
92	30	50.8	455	4	US-09-543-681A-8288	Sequence 8288, Ap
93	30	50.8	456	4	US-09-240-639-2	Sequence 2, Appli
94	30	50.8	468	4	US-09-543-681A-8164	Sequence 8164, Ap
95	30	50.8	471	4	US-09-608-285A-60	Sequence 60, Appl
96	30	50.8	473	4	US-09-252-991A-29434	Sequence 29434, A
97	30	50.8	473	4	US-09-134-000C-5487	Sequence 5487, Ap
98	30	50.8	479	4	US-09-328-352-4361	Sequence 4361, Ap
99	30	50.8	484	4	US-09-608-285A-27	Sequence 27, Appl
100	30	50.8	484	4	US-09-370-265-27	Sequence 27, Appl
101	30	50.8	484	4	US-09-557-800C-27	Sequence 27, Appl
102	30	50.8	484	4	US-09-370-625A-27	Sequence 27, Appl
103	30	50.8	487	4	US-09-107-532A-5731	Sequence 5731, Ap
104	30	50.8	491	4	US-09-543-681A-8202	Sequence 8202, Ap
105	30	50.8	510	4	US-09-489-039A-11778	Sequence 11778, A
106	30	50.8	529	4	US-09-711-164-431	Sequence 431, App
107	30	50.8	541	4	US-09-252-991A-31415	Sequence 31415, A
108	30	50.8	558	4	US-09-252-991A-18187	Sequence 18187, A
109	30	50.8	618	4	US-09-555-820A-14	Sequence 14, Appl
110	30	50.8	626	4	US-09-214-619-4	Sequence 4, Appli
111	30	50.8	626	4	US-09-555-820A-12	Sequence 12, Appl
112	30	50.8	634	4	US-09-252-991A-24131	Sequence 24131, A
113	30	50.8	646	4	US-09-489-039A-12750	Sequence 12750, A
114	30	50.8	665	2	US-08-846-762-17	Sequence 17, Appl
115	30	50.8	665	2	US-08-846-762-94	Sequence 94, Appl
116	30	50.8	671	4	US-09-252-991A-28798	Sequence 28798, A
117	30	50.8	703	4	US-09-328-352-7467	Sequence 7467, Ap
118	30	50.8	892	4	US-09-489-039A-14119	Sequence 14119, A
119	30	50.8	1843	3	US-09-413-814-50	Sequence 50, Appl
120	30	50.8	1882	4	US-09-369-364A-13	Sequence 13, Appl
121	29	49.2	30	2	US-08-404-531B-37	Sequence 37, Appl
122	29	49.2	30	3	US-08-476-900A-37	Sequence 37, Appl
123	29	49.2	30	3	US-08-488-546A-37	Sequence 37, Appl
124	29	49.2	35	3	US-08-942-686-5	Sequence 5, Appli
125	29	49.2	69	4	US-08-311-731A-340	Sequence 340, App



126	29	49.2	72	4	US-09-464-535-14	Sequence 14, Appl
127	29	49.2	89	4	US-09-464-535-40	Sequence 40, Appl
128	29	49.2	105	4	US-09-489-039A-7456	Sequence 7456, Ap
129	29	49.2	132	3	US-09-073-297-16	Sequence 16, Appl
130	29	49.2	148	2	US-08-392-625-18	Sequence 18, Appl
131	29	49.2	148	2	US-08-466-961A-18	Sequence 18, Appl
132	29	49.2	148	2	US-08-645-193B-13	Sequence 13, Appl
133	29	49.2	159	4	US-09-149-476-499	Sequence 499, App
134	29	49.2	197	4	US-09-813-918-4	Sequence 4, Appli
135	29	49.2	200	4	US-09-107-532A-5545	Sequence 5545, Ap
136	29	49.2	243	4	US-09-173-300-22	Sequence 22, Appl
137	29	49.2	273	4	US-09-252-991A-20770	Sequence 20770, A
138	29	49.2	277	4	US-09-489-039A-10481	Sequence 10481, A
139	29	49.2	292	4	US-09-328-352-4894	Sequence 4894, Ap
140	29	49.2	294	4	US-09-252-991A-22068	Sequence 22068, A
141	29	49.2	357	4	US-09-556-601-26	Sequence 26, Appl
142	29	49.2	365	4	US-09-252-991A-19530	Sequence 19530, A
143	29	49.2	400	4	US-09-252-991A-20800	Sequence 20800, A
144	29	49.2	409	4	US-09-540-236-2952	Sequence 2952, Ap
145	29	49.2	415	3	US-09-198-956-6	Sequence 6, Appli
146	29	49.2	415	4	US-09-670-141-6	Sequence 6, Appli
147	29	49.2	424	1	US-08-419-414-2	Sequence 2, Appli
148	29	49.2	434	2	US-08-468-812-6	Sequence 6, Appli
149	29	49.2	434	4	US-08-590-563-6	Sequence 6, Appli
150	29	49.2	434	4	US-09-770-621-6	Sequence 6, Appli

#### ALIGNMENTS

#### RESULT 1

US-08-313-288B-20

; Sequence 20, Application US/08313288B

; Patent No. 5750502

; GENERAL INFORMATION:

; APPLICANT: Jessell, Thomas M. and Avihu Klar

; TITLE OF INVENTION: CLONING, EXPRESSION AND USES OF A

; TITLE OF INVENTION: NOVEL SECRETED PROTEIN, F-SPONDIN

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/313,288B

; FILING DATE: January 5, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 40028-A-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0526  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1170 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-313-288B-20

Query Match 93.2%; Score 55; DB 1; Length 1170;  
Best Local Similarity 100.0%; Pred. No. 0.022;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVLQNVRFVF 12  
|||  
Db 209 QGVLQNVRFVF 219

#### RESULT 2

US-09-199-637A-267

; Sequence 267, Application US/09199637A  
; Patent No. 6355411  
; GENERAL INFORMATION:  
; APPLICANT: Ausubel, Frederick  
; APPLICANT: Goodman, Howard M.  
; APPLICANT: Rahme, Laurence G.  
; APPLICANT: Mahajan-Miklos, Shalina  
; APPLICANT: Tan, Man-Wah  
; APPLICANT: Cao, Hui  
; APPLICANT: Drenkard, Eliana  
; APPLICANT: Tsongalis, John  
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID  
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF  
; FILE REFERENCE: 00786/361002  
; CURRENT APPLICATION NUMBER: US/09/199,637A  
; CURRENT FILING DATE: 1998-11-25  
; PRIOR APPLICATION NUMBER: 60/066,517  
; PRIOR FILING DATE: 1997-11-25  
; NUMBER OF SEQ ID NOS: 437  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 267  
; LENGTH: 446  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-199-637A-267

Query Match 62.7%; Score 37; DB 4; Length 446;  
Best Local Similarity 72.7%; Pred. No. 26;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 AQGVLQNVRFV 11

OM protein - protein search, using sw model

Run on: April 7, 2004, 20:55:10 ; Search time 33.1429 Seconds  
 (without alignments)  
 95.091 Million cell updates/sec

Title: US-10-030-735-25  
 Perfect score: 59  
 Sequence: 1 AQQVLQNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	ID		Description

1	55	93.2	240	12	US-10-419-462-40	Sequence 40, Appl
2	55	93.2	466	9	US-09-925-301-1047	Sequence 1047, Ap
3	55	93.2	831	12	US-09-939-853A-97	Sequence 97, Appl
4	55	93.2	831	12	US-09-939-853A-98	Sequence 98, Appl
5	55	93.2	1152	9	US-09-919-603-1	Sequence 1, Appli
6	55	93.2	1170	12	US-10-211-462-38	Sequence 38, Appl
7	55	93.2	1170	12	US-10-231-956A-482	Sequence 482, App
8	55	93.2	1170	12	US-10-419-462-38	Sequence 38, Appl
9	55	93.2	1170	14	US-10-020-141-12	Sequence 12, Appl
10	55	93.2	1170	14	US-10-017-721-2	Sequence 2, Appli
11	55	93.2	1170	14	US-10-021-660-114	Sequence 114, App
12	55	93.2	1170	14	US-10-008-093-2	Sequence 2, Appli
13	55	93.2	1170	15	US-10-295-027-1170	Sequence 1170, Ap
14	39	66.1	15	15	US-10-285-394-153	Sequence 153, App
15	39	66.1	3500	14	US-10-153-219-2	Sequence 2, Appli
16	39	66.1	3537	14	US-10-153-219-15	Sequence 15, Appl
17	37	62.7	16	9	US-09-822-682-6	Sequence 6, Appli
18	37	62.7	151	9	US-09-925-299-1251	Sequence 1251, Ap
19	37	62.7	151	10	US-09-925-299-1251	Sequence 1251, Ap
20	37	62.7	446	10	US-09-975-719-267	Sequence 267, App
21	37	62.7	710	12	US-10-282-122A-75098	Sequence 75098, A
22	37	62.7	1168	9	US-09-919-603-2	Sequence 2, Appli
23	37	62.7	1172	9	US-09-919-770-4	Sequence 4, Appli
24	37	62.7	1172	9	US-09-822-682-2	Sequence 2, Appli
25	37	62.7	1172	10	US-09-866-050A-658	Sequence 658, App
26	37	62.7	1172	14	US-10-060-036-171	Sequence 171, App
27	37	62.7	1172	14	US-10-020-141-14	Sequence 14, Appl
28	37	62.7	1172	14	US-10-017-724-2	Sequence 2, Appli
29	37	62.7	1172	14	US-10-301-822-203	Sequence 203, App
30	37	62.7	1172	14	US-10-008-093-4	Sequence 4, Appli
31	37	62.7	1172	14	US-10-007-781-2	Sequence 2, Appli
32	37	62.7	1172	15	US-10-372-683-16	Sequence 16, Appl
33	37	62.7	1457	9	US-09-772-316-2	Sequence 2, Appli
34	37	62.7	1457	15	US-10-600-058-2	Sequence 2, Appli
35	36	61.0	75	12	US-10-424-599-283340	Sequence 283340,
36	36	61.0	87	12	US-10-424-599-254511	Sequence 254511,
37	36	61.0	410	15	US-10-369-493-8711	Sequence 8711, Ap
38	35	59.3	52	12	US-10-424-599-268809	Sequence 268809,
39	35	59.3	326	12	US-10-424-599-191204	Sequence 191204,
40	35	59.3	461	15	US-10-369-493-8924	Sequence 8924, Ap
41	35	59.3	498	12	US-10-425-114-49331	Sequence 49331, A
42	35	59.3	920	15	US-10-369-493-3545	Sequence 3545, Ap
43	35	59.3	1508	15	US-10-369-493-4104	Sequence 4104, Ap
44	34	57.6	51	12	US-10-424-599-193989	Sequence 193989,
45	34	57.6	51	12	US-10-424-599-198354	Sequence 198354,
46	34	57.6	72	12	US-10-424-599-147545	Sequence 147545,
47	34	57.6	215	14	US-10-029-386-33750	Sequence 33750, A
48	34	57.6	226	9	US-09-925-300-1217	Sequence 1217, Ap
49	34	57.6	263	15	US-10-369-493-4853	Sequence 4853, Ap
50	34	57.6	263	15	US-10-369-493-7612	Sequence 7612, Ap
51	34	57.6	265	12	US-10-282-122A-47647	Sequence 47647, A
52	34	57.6	277	9	US-09-738-626-4555	Sequence 4555, Ap
53	34	57.6	311	12	US-10-282-122A-70382	Sequence 70382, A
54	34	57.6	312	9	US-09-815-242-5267	Sequence 5267, Ap
55	34	57.6	329	12	US-10-424-599-177504	Sequence 177504,
56	34	57.6	333	15	US-10-369-493-5588	Sequence 5588, Ap

57	34	57.6	390	14	US-10-420-845-6	Sequence 6, Appli
58	34	57.6	426	12	US-10-282-122A-52492	Sequence 52492, A
59	34	57.6	466	12	US-10-424-599-232021	Sequence 232021,
60	34	57.6	537	12	US-10-425-114-49306	Sequence 49306, A
61	34	57.6	559	14	US-10-284-499-47	Sequence 47, Appl
62	34	57.6	597	12	US-10-282-122A-76883	Sequence 76883, A
63	34	57.6	651	14	US-10-420-845-5	Sequence 5, Appli
64	34	57.6	1070	14	US-10-420-845-22	Sequence 22, Appl
65	34	57.6	1504	9	US-09-932-145-7	Sequence 7, Appli
66	34	57.6	1504	14	US-10-284-499-1	Sequence 1, Appli
67	34	57.6	1504	15	US-10-395-812-7	Sequence 7, Appli
68	34	57.6	3063	14	US-10-177-293-61	Sequence 61, Appl
69	34	57.6	3063	14	US-10-177-293-63	Sequence 63, Appl
70	34	57.6	3063	14	US-10-301-822-26	Sequence 26, Appl
71	33	55.9	40	10	US-09-789-996-3	Sequence 3, Appli
72	33	55.9	50	11	US-09-833-245-167	Sequence 167, App
73	33	55.9	50	11	US-09-833-245-171	Sequence 171, App
74	33	55.9	81	12	US-10-424-599-153877	Sequence 153877,
75	33	55.9	110	9	US-09-925-297-556	Sequence 556, App
76	33	55.9	151	12	US-10-424-599-197953	Sequence 197953,
77	33	55.9	157	10	US-09-952-680A-33	Sequence 33, Appl
78	33	55.9	226	9	US-09-895-913A-362	Sequence 362, App
79	33	55.9	228	12	US-10-282-122A-49254	Sequence 49254, A
80	33	55.9	247	9	US-09-815-242-11342	Sequence 11342, A
81	33	55.9	247	9	US-09-815-242-11514	Sequence 11514, A
82	33	55.9	247	12	US-10-282-122A-58737	Sequence 58737, A
83	33	55.9	263	15	US-10-369-493-11024	Sequence 11024, A
84	33	55.9	315	12	US-10-282-122A-70417	Sequence 70417, A
85	33	55.9	350	10	US-09-789-996-23	Sequence 23, Appl
86	33	55.9	350	10	US-09-952-680A-24	Sequence 24, Appl
87	33	55.9	350	15	US-10-352-843-15	Sequence 15, Appl
88	33	55.9	354	10	US-09-789-996-21	Sequence 21, Appl
89	33	55.9	354	10	US-09-789-996-22	Sequence 22, Appl
90	33	55.9	354	10	US-09-789-996-24	Sequence 24, Appl
91	33	55.9	354	10	US-09-952-680A-25	Sequence 25, Appl
92	33	55.9	357	9	US-09-984-292-7	Sequence 7, Appli
93	33	55.9	357	9	US-09-989-497-7	Sequence 7, Appli
94	33	55.9	381	10	US-09-952-680A-29	Sequence 29, Appl
95	33	55.9	430	10	US-09-769-736-6	Sequence 6, Appli
96	33	55.9	432	12	US-10-424-599-228725	Sequence 228725,
97	32	54.2	45	12	US-10-424-599-275218	Sequence 275218,
98	32	54.2	53	12	US-10-424-599-276257	Sequence 276257,
99	32	54.2	75	12	US-10-424-599-165575	Sequence 165575,
100	32	54.2	81	14	US-10-029-386-29347	Sequence 29347, A
101	32	54.2	101	12	US-10-424-599-161904	Sequence 161904,
102	32	54.2	106	12	US-10-424-599-242530	Sequence 242530,
103	32	54.2	116	12	US-10-424-599-154870	Sequence 154870,
104	32	54.2	124	11	US-09-833-245-1905	Sequence 1905, Ap
105	32	54.2	125	12	US-10-424-599-259520	Sequence 259520,
106	32	54.2	163	12	US-10-424-599-237046	Sequence 237046,
107	32	54.2	180	12	US-10-424-599-228708	Sequence 228708,
108	32	54.2	188	9	US-09-925-300-1176	Sequence 1176, Ap
109	32	54.2	199	12	US-10-282-122A-46461	Sequence 46461, A
110	32	54.2	211	12	US-10-335-977-7519	Sequence 7519, Ap
111	32	54.2	224	12	US-10-335-977-7521	Sequence 7521, Ap
112	32	54.2	240	12	US-10-425-114-41336	Sequence 41336, A
113	32	54.2	297	15	US-10-108-260A-4132	Sequence 4132, Ap

114	32	54.2	331	9	US-09-738-626-4595	Sequence 4595, Ap
115	32	54.2	353	9	US-09-899-295-4	Sequence 4, Appli
116	32	54.2	353	9	US-09-899-295-6	Sequence 6, Appli
117	32	54.2	353	9	US-09-899-295-8	Sequence 8, Appli
118	32	54.2	353	9	US-09-984-292-2	Sequence 2, Appli
119	32	54.2	353	9	US-09-984-292-4	Sequence 4, Appli
120	32	54.2	353	9	US-09-984-292-5	Sequence 5, Appli
121	32	54.2	353	9	US-09-984-292-6	Sequence 6, Appli
122	32	54.2	353	9	US-09-984-292-8	Sequence 8, Appli
123	32	54.2	353	9	US-09-984-292-14	Sequence 14, Appl
124	32	54.2	353	9	US-09-984-292-16	Sequence 16, Appl
125	32	54.2	353	9	US-09-984-292-17	Sequence 17, Appl
126	32	54.2	353	9	US-09-984-292-18	Sequence 18, Appl
127	32	54.2	353	9	US-09-984-292-19	Sequence 19, Appl
128	32	54.2	353	9	US-09-984-292-24	Sequence 24, Appl
129	32	54.2	353	9	US-09-989-497-2	Sequence 2, Appli
130	32	54.2	353	9	US-09-989-497-4	Sequence 4, Appli
131	32	54.2	353	9	US-09-989-497-5	Sequence 5, Appli
132	32	54.2	353	9	US-09-989-497-6	Sequence 6, Appli
133	32	54.2	353	9	US-09-989-497-8	Sequence 8, Appli
134	32	54.2	353	9	US-09-989-497-14	Sequence 14, Appl
135	32	54.2	353	9	US-09-989-497-16	Sequence 16, Appl
136	32	54.2	353	9	US-09-989-497-17	Sequence 17, Appl
137	32	54.2	353	9	US-09-989-497-18	Sequence 18, Appl
138	32	54.2	353	9	US-09-989-497-19	Sequence 19, Appl
139	32	54.2	353	9	US-09-989-497-24	Sequence 24, Appl
140	32	54.2	353	10	US-09-952-680A-27	Sequence 27, Appl
141	32	54.2	353	11	US-09-899-295-4	Sequence 4, Appli
142	32	54.2	353	11	US-09-899-295-6	Sequence 6, Appli
143	32	54.2	353	11	US-09-899-295-8	Sequence 8, Appli
144	32	54.2	353	15	US-10-352-843-13	Sequence 13, Appl
145	32	54.2	355	15	US-10-369-493-4982	Sequence 4982, Ap
146	32	54.2	355	15	US-10-352-843-12	Sequence 12, Appl
147	32	54.2	359	9	US-09-899-295-2	Sequence 2, Appli
148	32	54.2	359	9	US-09-984-292-1	Sequence 1, Appli
149	32	54.2	359	9	US-09-984-292-3	Sequence 3, Appli
150	32	54.2	359	9	US-09-984-292-9	Sequence 9, Appli

#### ALIGNMENTS

##### RESULT 1

US-10-419-462-40

; Sequence 40, Application US/10419462

; Publication No. US20040053392A1

; GENERAL INFORMATION:

; APPLICANT: Kevin J. Williams

; APPLICANT: Williams, Kevin J.

; TITLE OF INVENTION: Thrombospondin Fragments and Uses Thereof In Clinical Assays for

; TITLE OF INVENTION: Cancer and Generation of Antibodies and Other Binding Agents

; FILE REFERENCE: W1107-20005

; CURRENT APPLICATION NUMBER: US/10/419,462

; CURRENT FILING DATE: 2003-04-17

; NUMBER OF SEQ ID NOS: 53

; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 40  
; LENGTH: 240  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Thrombospondin Region plus N-terminal domain  
US-10-419-462-40

Query Match 93.2%; Score 55; DB 12; Length 240;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVLQNVRFVF 12  
|||  
Db 191 QGVLQNVRFVF 201

#### RESULT 2

US-09-925-301-1047  
; Sequence 1047, Application US/09925301  
; Patent No. US20020052308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA106  
; CURRENT APPLICATION NUMBER: US/09/925,301  
; CURRENT FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05882  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1694  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1047  
; LENGTH: 466  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-925-301-1047

Query Match 93.2%; Score 55; DB 9; Length 466;  
Best Local Similarity 100.0%; Pred. No. 0.03;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 QGVLQNVRFVF 12  
|||  
Db 262 QGVLQNVRFVF 272

#### RESULT 3

US-09-939-853A-97  
; Sequence 97, Application US/09939853A  
; Publication No. US20040039163A1  
; GENERAL INFORMATION:  
; APPLICANT: Burgess et al.  
; TITLE OF INVENTION: No. US20040039163A1el Proteins and Nucleic Acids Encoding  
Same

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:36:09 ; Search time 10.4416 Seconds  
 (without alignments)  
 110.548 Million cell updates/sec

Title: US-10-030-735-25  
 Perfect score: 59  
 Sequence: 1 AQGVLQNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	55	93.2	229	2	S57957	thrombospondin 1 -
2	55	93.2	1170	1	TSHUP1	thrombospondin 1 p
3	55	93.2	1170	2	A40558	thrombospondin 1 p
4	37	62.7	151	2	C57253	tRNA-pseudouridine
5	37	62.7	1172	1	TSHUP2	thrombospondin 2 p
6	37	62.7	1172	2	A42587	thrombospondin 2 p
7	36	61.0	186	2	B75421	probable pilin, ty
8	35	59.3	40	2	T08870	alternative respir
9	35	59.3	89	2	B39529	cadherin-associate
10	35	59.3	102	2	D81910	hypothetical prote
11	35	59.3	326	2	T08849	alternative respir
12	35	59.3	747	2	AE2929	two component resp
13	35	59.3	783	2	A98353	probable transcrip



14	35	59.3	927	2	T43110	lactacin 481/lacto
15	35	59.3	1178	1	A39804	thrombospondin pre
16	34	57.6	102	2	G81105	hypothetical prote
17	34	57.6	106	2	C90261	hypothetical prote
18	34	57.6	265	2	T40878	probable FAD synth
19	34	57.6	308	2	JC5468	leukocidin chain 1
20	34	57.6	311	2	C89968	leukotoxin LukeE [i
21	34	57.6	333	2	JC2432	L-lactate dehydrog
22	34	57.6	337	2	E97882	hypothetical prote
23	34	57.6	417	2	H83708	hypothetical prote
24	34	57.6	595	2	A43534	Lupus autoantigen
25	34	57.6	597	2	B82881	hypothetical prote
26	34	57.6	993	2	T17230	hypothetical prote
27	34	57.6	1308	2	T05178	hypothetical prote
28	33	55.9	54	2	S35697	leukocidin chain F
29	33	55.9	175	2	S75258	hypothetical prote
30	33	55.9	212	2	D81929	probable imidazole
31	33	55.9	247	1	A64590	probable 3-oxoacyl
32	33	55.9	247	2	B71923	3-oxoacyl-[acyl-ca
33	33	55.9	286	2	C49238	gamma-hemolysin co
34	33	55.9	298	2	C87403	FdhD protein [impo
35	33	55.9	310	2	S68225	synergohymenotropi
36	33	55.9	312	2	T00160	leukocidin chain S
37	33	55.9	312	2	S32211	leucocidin chain S
38	33	55.9	315	2	PC4078	hlgC-like protein
39	33	55.9	315	2	A49234	leucocidin R S com
40	33	55.9	315	2	JN0626	leukocidin chain S
41	33	55.9	315	2	E90043	gamma-hemolysin co
42	33	55.9	326	2	S54267	repA protein - Bac
43	33	55.9	331	2	S54263	rep A protein - Ba
44	33	55.9	350	1	RGHUT1	GTP-binding regula
45	33	55.9	350	1	RGBOT1	GTP-binding regula
46	33	55.9	350	1	RGMST1	GTP-binding regula
47	33	55.9	354	1	RGHUT2	GTP-binding regula
48	33	55.9	354	1	RGBOT2	GTP-binding regula
49	33	55.9	354	2	S24352	gustducin - rat
50	33	55.9	363	2	T50582	FecCD-family membr
51	33	55.9	381	2	A48071	guanine nucleotide
52	33	55.9	405	2	G82165	glucose-1-phosphat
53	33	55.9	455	2	D83264	hypothetical prote
54	33	55.9	511	2	S77350	hypothetical prote
55	33	55.9	780	2	T50315	hypothetical prote
56	33	55.9	1307	2	T30882	multidrug resistan
57	32	54.2	126	2	T20427	hypothetical prote
58	32	54.2	145	2	T20985	hypothetical prote
59	32	54.2	162	2	B88349	protein F15D4.3 [i
60	32	54.2	194	2	T04180	ribosomal protein
61	32	54.2	207	2	S53801	chitin synthase (E
62	32	54.2	207	2	T48527	hypothetical prote
63	32	54.2	212	2	AE2158	hypothetical prote
64	32	54.2	226	2	AI2668	hypothetical prote
65	32	54.2	267	2	G97450	ABC transporter, A
66	32	54.2	300	2	D81399	malate dehydrogena
67	32	54.2	332	2	I67791	cytochrome P450 2B
68	32	54.2	353	2	S34347	GTP-binding regula
69	32	54.2	353	2	B40891	GTP-binding protei
70	32	54.2	354	2	S33309	GTP-binding regula

71	32	54.2	355	2	A40891	GTP-binding protei
72	32	54.2	355	2	A41534	GTP-binding protei
73	32	54.2	355	2	T15288	hypothetical prote
74	32	54.2	359	1	RGHUGY	GTP-binding regula
75	32	54.2	359	1	RGMSQ	GTP-binding regula
76	32	54.2	359	1	RGMS11	GTP-binding regula
77	32	54.2	359	2	S45699	GTP-binding regula
78	32	54.2	359	2	S71963	GTP-binding protei
79	32	54.2	359	2	S45700	G-alpha-11 protein
80	32	54.2	359	2	S30359	GTP-binding regula
81	32	54.2	371	2	B69451	conserved hypothet
82	32	54.2	378	2	A98188	probable membrane
83	32	54.2	378	2	A86035	probable membrane
84	32	54.2	378	2	S47818	probable membrane
85	32	54.2	379	2	A41095	GTP-binding regula
86	32	54.2	382	1	I39848	cell division init
87	32	54.2	382	2	F83969	cell-division init
88	32	54.2	390	1	S58814	cell division prot
89	32	54.2	392	2	G69896	conserved hypothet
90	32	54.2	397	2	G64703	type IIS restricti
91	32	54.2	406	2	G71816	probable type II r
92	32	54.2	415	2	G90705	probable transport
93	32	54.2	415	2	B85556	probable transport
94	32	54.2	415	2	G64790	ybdG protein - Esc
95	32	54.2	425	2	D82684	UDP-N-acetylglucos
96	32	54.2	434	2	D72353	lipopolysaccharide
97	32	54.2	440	2	F96556	IAA-Ala hydrolase
98	32	54.2	448	2	D87263	hypothetical prote
99	32	54.2	452	2	F91292	hypothetical prote
100	32	54.2	452	2	A86134	hypothetical prote
101	32	54.2	453	2	H64974	hypothetical prote
102	32	54.2	453	2	A98990	hypothetical prote
103	32	54.2	453	2	D85835	hypothetical prote
104	32	54.2	453	2	AE0774	probable protease
105	32	54.2	454	2	D86793	drug-export protei
106	32	54.2	454	2	T03130	probable tyrosine
107	32	54.2	456	2	H97735	hypothetical prote
108	32	54.2	464	2	AG0347	probable proteinas
109	32	54.2	467	2	D84938	H+-transporting tw
110	32	54.2	483	2	A87583	peptidoglycan bind
111	32	54.2	487	1	H81109	IMP dehydrogenase
112	32	54.2	487	2	F81906	IMP dehydrogenase
113	32	54.2	492	2	S32491	testosterone 7alph
114	32	54.2	492	2	A31887	testosterone 7alph
115	32	54.2	492	2	A34272	testosterone 7alph
116	32	54.2	558	2	T48150	stress-induced pro
117	32	54.2	562	2	AD0328	probable potassium
118	32	54.2	584	2	T29469	hypothetical prote
119	32	54.2	628	2	AF1108	transporter homolo
120	32	54.2	628	2	AG1469	transporter homolo
121	32	54.2	642	2	T51421	L-aspartate oxidas
122	32	54.2	681	2	T16353	hypothetical prote
123	32	54.2	838	2	T47828	hypothetical prote
124	32	54.2	855	2	C90262	aconitate hydratas
125	32	54.2	889	2	JC6015	chitin synthase (E
126	32	54.2	906	2	G70767	probable helicase
127	32	54.2	920	2	G87075	probable helicase,

128	32	54.2	1251	2	C82721	conserved hypothet
129	32	54.2	1362	2	T41534	leptomycin B resis
130	32	54.2	1413	2	B82877	conserved hypothet
131	32	54.2	1483	2	C97012	probably celluloso
132	32	54.2	1835	2	S46082	urea carboxylase (
133	32	54.2	4572	2	S57908	hypothetical 527K
134	32	54.2	6658	2	T13931	projectin - fruit
135	31.5	53.4	387	1	TVECG	phosphoglycerate k
136	31.5	53.4	387	2	E91103	phosphoglycerate k
137	31.5	53.4	387	2	H85948	phosphoglycerate k
138	31.5	53.4	387	2	AD0875	phosphoglycerate k
139	31	52.5	96	2	S47297	suilysin - Strepto
140	31	52.5	102	2	G84013	hypothetical prote
141	31	52.5	112	2	E83817	hypothetical prote
142	31	52.5	152	2	S44740	C02C2.2 protein -
143	31	52.5	156	2	S60953	iron-sulfur cofact
144	31	52.5	189	2	B45190	chitin synthase (E
145	31	52.5	195	2	H45189	chitin synthase (E
146	31	52.5	198	2	A45190	chitin synthase (E
147	31	52.5	198	2	G45189	chitin synthase (E
148	31	52.5	217	2	H64419	transaldolase (EC
149	31	52.5	218	2	G72394	transaldolase-rela
150	31	52.5	234	2	T11914	cox1 intron protei

#### ALIGNMENTS

##### RESULT 1

S57957

thrombospondin 1 - bovine (fragment)

C;Species: Bos primigenius taurus (cattle)

C;Date: 13-Jan-1996 #sequence\_revision 19-Apr-1996 #text\_change 20-Aug-1999

C;Accession: S57957

R;Lafeuillade, B.; Pellerin, S.; Keramidas, M.; Chambaz, E.M.; Feige, J.J.

submitted to the EMBL Data Library, July 1995

A;Description: Opposite regulation of thrombospondin-1 and CISP/thrombospondin-2 expression by ACTH in adrenocortical cells.

A;Reference number: S57955

A;Accession: S57957

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-229 <LAF>

A;Cross-references: EMBL:X89511; NID:g899228; PIDN:CAA61682.1; PID:g899229

C;Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von Willebrand factor type C repeat homology

Query Match 93.2%; Score 55; DB 2; Length 229;

Best Local Similarity 100.0%; Pred. No. 0.0015;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      2 QGVLQNVRFVF 12
          |||||
Db     191 QGVLQNVRFVF 201

```

##### RESULT 2

TSHUP1

thrombospondin 1 precursor - human

C;Species: Homo sapiens (man)

C;Date: 23-Aug-1987 #sequence\_revision 03-Aug-1995 #text\_change 17-Nov-2000

C;Accession: A26155; A34274; A30140; A25812; A05172; A42927

R;Lawler, J.; Hynes, R.O.

J. Cell Biol. 103, 1635-1648, 1986

A;Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins.

A;Reference number: A26155; MUID:87057617; PMID:2430973

A;Accession: A26155

A;Molecule type: mRNA

A;Residues: 1-1170 <LAW>

A;Cross-references: GB:X04665; NID:g37137; PIDN:CAA28370.1; PID:g37138

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Laherty, C.D.; Gierman, T.M.; Dixit, V.M.

J. Biol. Chem. 264, 11222-11227, 1989

A;Title: Characterization of the promoter region of the human thrombospondin gene. DNA sequences within the first intron increase transcription.

A;Reference number: A34274; MUID:89291870; PMID:2544587

A;Accession: A34274

A;Molecule type: DNA

A;Residues: 1-166 <LAH>

A;Cross-references: GB:J04835

R;Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein, P.; Frazier, W.A.

J. Cell Biol. 108, 729-736, 1989

A;Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in the 3' untranslated region.

A;Reference number: A30140; MUID:89139590; PMID:2918029

A;Accession: A30140

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-522,'A',524-1170 <HEN>

A;Cross-references: EMBL:X14787; NID:g37464; PIDN:CAA32889.1; PID:g37465

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.

Biochemistry 25, 8418-8425, 1986

A;Title: Partial amino acid sequence of human thrombospondin as determined by analysis of cDNA clones: homology to malarial circumsporozoite proteins.

A;Reference number: A25812; MUID:87157592; PMID:3030396

A;Accession: A25812

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-397 <KOB>

A;Cross-references: GB:M25631; NID:g538353; PIDN:AAA36741.1; PID:g538354

R;Dixit, V.M.; Hennessy, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A;Reference number: A05172; MUID:86287276; PMID:3461443

A;Accession: A05172

A;Molecule type: mRNA

A;Residues: 1-83,'A',85-374,'RC' <DIX>

A;Cross-references: GB:M14326; NID:g340005; PIDN:AAA61237.1; PID:g553801

A;Note: parts of this sequence, including the amino end of the mature protein, were determined by protein sequencing

R;Sun, X.; Skorstengaard, K.; Mosher, D.F.

J. Cell Biol. 118, 693-701, 1992

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:22:56 ; Search time 5.61039 Seconds  
 (without alignments)  
 111.372 Million cell updates/sec

Title: US-10-030-735-25  
 Perfect score: 59  
 Sequence: 1 AQQVLQNVRFVF 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	55	93.2	1170	1	TSP1_BOVIN	Q28178 bos taurus
2	55	93.2	1170	1	TSP1_HUMAN	P07996 homo sapien
3	55	93.2	1170	1	TSP1_MOUSE	P35441 mus musculu
4	55	93.2	1173	1	TSP1_XENLA	P35448 xenopus lae
5	39	66.1	259	1	TRMB_AZOSE	Q8g9c6 azoarcus sp
6	37	62.7	427	1	XRCL_PSEAE	Q9f771 pseudomonas
7	37	62.7	1172	1	TSP2_HUMAN	P35442 homo sapien
8	37	62.7	1172	1	TSP2_MOUSE	Q03350 mus musculu
9	35	59.3	326	1	AOX3_SOYBN	O03376 glycine max
10	35	59.3	1170	1	TSP2_BOVIN	Q95116 bos taurus
11	35	59.3	1178	1	TSP2_CHICK	P35440 gallus gall
12	34	57.6	250	1	UBIE_COXBU	Q83a90 coxiella bu
13	34	57.6	265	1	FAD1_SCHPO	O74841 schizosacch
14	34	57.6	333	1	LDH_CAEEL	Q27888 caenorhabdi
15	34	57.6	608	1	KU70_MOUSE	P23475 mus musculu
16	34	57.6	700	1	V018_FOWPV	Q9j5i3 fowlpox vir
17	34	57.6	3063	1	CA1C_HUMAN	Q99715 homo sapien

18	34	57.6	3119	1	CA1C_MOUSE	Q60847	mus musculu
19	33	55.9	212	1	HIS5_NEIMA	Q9jvh3	neisseria m
20	33	55.9	315	1	HLGC_STAAU	Q07227	staphylococ
21	33	55.9	315	1	LUKS_STAAU	P31716	staphylococ
22	33	55.9	349	1	GBT1_BOVIN	P04695	bos taurus
23	33	55.9	349	1	GBT1_CANFA	Q28300	canis famil
24	33	55.9	349	1	GBT1_HUMAN	P11488	homo sapien
25	33	55.9	349	1	GBT1_MOUSE	P20612	mus musculu
26	33	55.9	353	1	GBT2_BOVIN	P04696	bos taurus
27	33	55.9	353	1	GBT2_HUMAN	P19087	homo sapien
28	33	55.9	353	1	GBT2_MOUSE	P50149	mus musculu
29	33	55.9	353	1	GBT3_RAT	P29348	rattus norv
30	33	55.9	380	1	GB12_HUMAN	Q03113	homo sapien
31	33	55.9	405	1	GLC1_VIBCH	Q9krb5	vibrio chol
32	32	54.2	126	1	Y334_BUCBP	Q89ag0	buchnera ap
33	32	54.2	353	1	GBQ_CANFA	Q28294	canis famil
34	32	54.2	353	1	GBQ_HOMAM	P91950	homarus ame
35	32	54.2	353	1	GBQ_HUMAN	P50148	homo sapien
36	32	54.2	353	1	GBQ_LYMST	P38411	lymnaea sta
37	32	54.2	353	1	GBQ_MOUSE	P21279	mus musculu
38	32	54.2	353	1	GBQ_PATYE	O15975	patinopecte
39	32	54.2	353	1	GBQ_RAT	P82471	rattus norv
40	32	54.2	353	1	GBQ_XENLA	P38410	xenopus lae
41	32	54.2	354	1	GB14_XENLA	O73819	xenopus lae
42	32	54.2	354	1	GBQ_LOLFO	P38412	loligo forb
43	32	54.2	355	1	GB14_BOVIN	P38408	bos taurus
44	32	54.2	355	1	GB14_MOUSE	P30677	mus musculu
45	32	54.2	359	1	GB11_BOVIN	P38409	bos taurus
46	32	54.2	359	1	GB11_HUMAN	P29992	homo sapien
47	32	54.2	359	1	GB11_MELGA	P45645	meleagris g
48	32	54.2	359	1	GB11_MOUSE	P21278	mus musculu
49	32	54.2	359	1	GB11_RAT	Q9jid2	rattus norv
50	32	54.2	359	1	GB11_XENLA	P43444	xenopus lae
51	32	54.2	364	1	NK23_HUMAN	Q8tau0	homo sapien
52	32	54.2	378	1	GB12_MOUSE	P27600	mus musculu
53	32	54.2	378	1	GB12_RAT	Q63210	rattus norv
54	32	54.2	378	1	YIBH_ECOLI	P32107	escherichia
55	32	54.2	382	1	FTSZ_BACHD	Q9k9t7	bacillus ha
56	32	54.2	382	1	FTSZ_BACSU	P17865	bacillus su
57	32	54.2	390	1	FTSZ_STAAM	P45498	staphylococ
58	32	54.2	392	1	OXDD_BACSU	O34767	bacillus su
59	32	54.2	394	1	FTSZ_STAEP	Q8cpk4	staphylococ
60	32	54.2	415	1	YBDG_ECOLI	P39455	escherichia
61	32	54.2	419	1	HPPD_MYCGR	O42764	mycosphaere
62	32	54.2	425	1	MURA_XYLFA	Q9pdg4	xylella fas
63	32	54.2	425	1	MURA_XYLFT	Q87dn8	xylella fas
64	32	54.2	453	1	YEGQ_ECOLI	P76403	escherichia
65	32	54.2	467	1	FLII_BUCAI	P57178	buchnera ap
66	32	54.2	492	1	CPA1_RAT	P11711	rattus norv
67	32	54.2	492	1	CPA2_RAT	P15149	rattus norv
68	32	54.2	492	1	CPAC_MOUSE	P56593	mus musculu
69	32	54.2	562	1	ATKA_YERPE	Q8zd96	yersinia pe
70	32	54.2	604	1	SR68_DROME	Q9vss2	drosophila
71	32	54.2	889	1	CHSC_ASPFU	Q92197	aspergillus
72	32	54.2	906	1	HELY_MYCTU	Q10701	mycobacteri
73	32	54.2	920	1	HELY_MYCLE	Q9zbd8	mycobacteri
74	32	54.2	1362	1	PMD1_SCHPO	P36619	schizosacch

75	32	54.2	1835	1	DUR1_YEAST	P32528	saccharomyc
76	31.5	53.4	386	1	PGK_ECO57	Q8xd03	escherichia
77	31.5	53.4	386	1	PGK_ECOLI	P11665	escherichia
78	31.5	53.4	386	1	PGK_SALTY	Q8xg18	salmonella
79	31	52.5	194	1	CHS2_AJECA	P30577	ajellomyces
80	31	52.5	198	1	CHS2_RHIAT	P30593	rhinocladie
81	31	52.5	217	1	TAL_METJA	Q58370	methanococc
82	31	52.5	218	1	TAL_THEMEA	Q9wyd1	thermotoga
83	31	52.5	298	1	RT03_ACACA	P46754	acanthamoeb
84	31	52.5	326	1	UL31_HSVEB	P28951	equine herp
85	31	52.5	335	1	Y286_STAEP	Q8ctu5	staphylococ
86	31	52.5	335	1	Y479_STAAM	Q8nxy9	staphylococ
87	31	52.5	335	1	Y524_STAAM	Q99w79	staphylococ
88	31	52.5	353	1	GB0_XENLA	P10825	xenopus lae
89	31	52.5	358	1	BUK_OCEIH	Q8cxe5	oceanobacil
90	31	52.5	380	1	PEX2_YARLI	Q99155	yarrowia li
91	31	52.5	382	1	YFPF_BACSU	P54166	bacillus su
92	31	52.5	384	1	GLGC_FUSNN	Q8rf63	fusobacteri
93	31	52.5	385	1	YFD0_YEAST	P43567	saccharomyc
94	31	52.5	386	1	PGK_PHOLU	Q8gf87	photorhabdu
95	31	52.5	386	1	PGK_VIBPA	Q87111	vibrio para
96	31	52.5	386	1	PGK_VIBVU	Q8dca0	vibrio vuln
97	31	52.5	387	1	PGK_VIBCH	P96154	vibrio chol
98	31	52.5	387	1	PGK_YERPE	Q8zhh3	yersinia pe
99	31	52.5	394	1	PGK_BACME	P24269	bacillus me
100	31	52.5	394	1	PGK_BACSU	P40924	bacillus su
101	31	52.5	394	1	PGK_WIGBR	Q8d2p9	wiggleswort
102	31	52.5	400	1	CPXP_RHISN	P55544	rhizobium s
103	31	52.5	406	1	T230_HUMAN	P48775	homo sapien
104	31	52.5	436	1	BRNQ_HAEIN	P71345	haemophilus
105	31	52.5	442	1	SSNA_ECOLI	Q46812	escherichia
106	31	52.5	487	1	IMDH_PASMU	Q916b7	pasteurella
107	31	52.5	488	1	IMDH_HAEIN	P44334	haemophilus
108	31	52.5	577	1	THT1_SCHPO	Q09684	schizosacch
109	31	52.5	585	1	YK05_CAEEL	P34293	caenorhabdi
110	31	52.5	680	1	DNK3_SYNEL	Q8dh10	synechococc
111	31	52.5	732	1	YM8K_YEAST	Q03254	saccharomyc
112	31	52.5	863	1	SIP1_YEAST	P32578	saccharomyc
113	31	52.5	885	1	CHS3_EXODE	P30602	exophiala d
114	31	52.5	911	1	CHSG_ASPFU	P54267	aspergillus
115	31	52.5	916	1	CHSB_EMENI	Q00757	emericella
116	31	52.5	975	1	CUT1_CANFA	P39881	canis famil
117	31	52.5	1045	1	RAG1_XENLA	Q91829	xenopus lae
118	31	52.5	1288	1	OPLA_MOUSE	Q8k010	mus musculu
119	31	52.5	1288	1	OPLA_RAT	P97608	rattus norv
120	31	52.5	1839	1	ALE3_AZОВI	Q44496	azotobacter
121	31	52.5	2167	1	BPH1_YEAST	P25356	saccharomyc
122	31	52.5	3124	1	CA1C_CHICK	P13944	gallus gall
123	30	50.8	70	1	Y414_HAEIN	Q57392	haemophilus
124	30	50.8	155	1	MLC1_DROPS	Q24621	drosophila
125	30	50.8	182	1	INB_MOUSE	P01575	mus musculu
126	30	50.8	196	1	BIM_MOUSE	O54918	mus musculu
127	30	50.8	196	1	BIM_RAT	O88498	rattus norv
128	30	50.8	199	1	HIS5_HAEIN	P44340	haemophilus
129	30	50.8	218	1	TAL2_LISMO	Q92eu7	listeria mo
130	30	50.8	222	1	PLL2_MOUSE	P09586	mus musculu
131	30	50.8	235	1	UL34_HSVSA	Q01045	herpesvirus

132	30	50.8	250	1	VGLL_HSV6U	P52508	human herpe
133	30	50.8	256	1	PQQC_RHIME	Q9exv0	rhizobium m
134	30	50.8	261	1	SMTA_ECOLI	P36566	escherichia
135	30	50.8	274	1	UPK_DEIRA	Q9rx61	deinococcus
136	30	50.8	291	1	CU59_DROME	Q9vzh1	drosophila
137	30	50.8	292	1	APAH_XYLFA	Q9pbj4	xylella fas
138	30	50.8	295	1	APAH_XYLFT	Q87c83	xylella fas
139	30	50.8	311	1	RLAO_CAEEL	Q93572	caenorhabdi
140	30	50.8	312	1	ANRA_MOUSE	Q99pe2	mus musculu
141	30	50.8	313	1	ANRA_HUMAN	Q9h9e1	homo sapien
142	30	50.8	349	1	GBT_XENLA	P38407	xenopus lae
143	30	50.8	353	1	RPOA_MYCSP	P38018	mycoplasma
144	30	50.8	355	1	GB14_HUMAN	Q95837	homo sapien
145	30	50.8	356	1	GBAG_CAEEL	Q9n2v6	caenorhabdi
146	30	50.8	358	1	GBA4_CAEEL	Q9big5	caenorhabdi
147	30	50.8	388	1	PGK_BUCAP	Q8k9b3	buchnera ap
148	30	50.8	393	1	PGK_BUCBP	P59461	buchnera ap
149	30	50.8	405	1	ARGJ_PSEAE	Q9hw04	pseudomonas
150	30	50.8	405	1	GLC1_VIBPA	Q87qx6	vibrio para

#### ALIGNMENTS

##### RESULT 1

##### TSP1\_BOVIN

ID TSP1\_BOVIN STANDARD; PRT; 1170 AA.  
AC Q28178; Q28179;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Thrombospondin 1 precursor.  
GN THBS1 OR TSP1 OR TSP-1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Holstein; TISSUE=Tooth;  
RX MEDLINE=98173773; PubMed=9507054;  
RA Ueno A., Yamashita K., Nagata T., Tsurumi C., Miwa Y., Kitamura S.,  
RA Inoue H.;  
RT "cDNA cloning of bovine thrombospondin 1 and its expression in  
RT odontoblasts and predentin.";  
RL Biochim. Biophys. Acta 1382:17-22(1998).  
RN [2]  
RP SEQUENCE OF 1-18 AND 710-1170 FROM N.A.  
RC TISSUE=Aortic endothelium;  
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;  
RL Submitted (JUN-1995) to the EMBL/GenBank/DDBJ databases.  
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-  
CC V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis  
CC and/or maintenance of dentin and dental pulp.



```

CC      -!- SUBUNIT: Homotrimer; disulfide-linked.
CC      -!- TISSUE SPECIFICITY: Odontoblasts.
CC      -!- SIMILARITY: Belongs to the thrombospondin family.
CC      -!- SIMILARITY: Contains 1 VWFC domain.
CC      -!- SIMILARITY: Contains 3 EGF-like domains.
CC      -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC      -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC      -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AB005287; BAA21115.1; -.
DR      EMBL; X87618; CAA60950.1; -.
DR      EMBL; X87619; CAA60951.1; -.
DR      PIR; S55501; S55501.
DR      GlycoSuiteDB; Q28178; -.
DR      InterPro; IPR001881; EGF_Ca.
DR      InterPro; IPR006209; EGF_like.
DR      InterPro; IPR006210; IEGF.
DR      InterPro; IPR000884; TSP1.
DR      InterPro; IPR008085; TSP_1.
DR      InterPro; IPR003367; tsp_3.
DR      InterPro; IPR008859; TSPC.
DR      InterPro; IPR003129; TSPN.
DR      InterPro; IPR001007; VWF_C.
DR      Pfam; PF00008; EGF; 2.
DR      Pfam; PF00090; tsp_1; 3.
DR      Pfam; PF02412; tsp_3; 13.
DR      Pfam; PF05735; TSPC; 1.
DR      Pfam; PF02210; TSPN; 1.
DR      Pfam; PF00093; vwc; 1.
DR      PRINTS; PR01705; TSP1REPEAT.
DR      SMART; SM00181; EGF; 3.
DR      SMART; SM00209; TSP1; 3.
DR      SMART; SM00210; TSPN; 1.
DR      SMART; SM00214; VWC; 1.
DR      PROSITE; PS00022; EGF_1; FALSE_NEG.
DR      PROSITE; PS01186; EGF_2; 1.
DR      PROSITE; PS50026; EGF_3; 2.
DR      PROSITE; PS50092; TSP1; 3.
DR      PROSITE; PS01208; VWFC_1; 1.
DR      PROSITE; PS50184; VWFC_2; 1.
KW      Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW      EGF-like domain; Signal.
FT      SIGNAL          1      18
FT      CHAIN           19    1170
FT      DOMAIN          19    232
FT      DOMAIN          24    221
FT      DOMAIN          316    373
FT      DOMAIN          379    429
FT      DOMAIN          435    490

```

			BY SIMILARITY.
			THROMBOSPONDIN 1.
			HEPARIN-BINDING (POTENTIAL).
			TSP N-TERMINAL.
			VWFC.
			TSP TYPE-1 1.
			TSP TYPE-1 2.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:33:49 ; Search time 31.2208 Seconds  
 (without alignments)  
 121.272 Million cell updates/sec

Title: US-10-030-735-25  
 Perfect score: 59  
 Sequence: 1 A Q G V L Q N V R F V F 12

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp Vertebrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	Match	Length	DB	ID	Description
-----							

1	55	93.2	229	6	Q28194	Q28194 bos taurus
2	55	93.2	496	13	Q7SY84	Q7sy84 xenopus lae
3	55	93.2	1171	11	Q8CGB2	Q8cgb2 mus musculu
4	55	93.2	1171	11	Q80YQ1	Q80yq1 mus musculu
5	39	66.1	593	2	Q8GH66	Q8gh66 mycobacteri
6	39	66.1	595	5	Q86NR6	Q86nr6 drosophila
7	39	66.1	721	5	Q9VTH0	Q9vth0 drosophila
8	39	66.1	1034	5	Q7Z2B9	Q7z2b9 trypanosoma
9	39	66.1	3242	9	Q859P9	Q859p9 bacterioph
10	38	64.4	101	16	Q7VC17	Q7vc17 prochloroco
11	37	62.7	210	2	Q9ZEN6	Q9zen6 wolinnella s
12	37	62.7	229	2	Q9EU71	Q9eu71 pseudomonas
13	37	62.7	426	2	Q7WZ26	Q7wz26 pseudomonas
14	37	62.7	427	2	Q7WY55	Q7wy55 pseudomonas
15	37	62.7	1172	11	Q8CG21	Q8cg21 mus musculu
16	37	62.7	1172	11	Q7TMT3	Q7tmt3 mus musculu
17	37	62.7	1457	12	Q9DY97	Q9dy97 porcine rep
18	37	62.7	1457	12	Q9WJB3	Q9wjb3 porcine rep
19	37	62.7	1457	12	Q9WBQ4	Q9wbq4 porcine rep
20	37	62.7	1457	12	Q91F53	Q91f53 porcine rep
21	37	62.7	1457	12	Q80KX0	Q80kx0 porcine rep
22	37	62.7	1460	12	Q8QQW9	Q8qqw9 porcine rep
23	37	62.7	1460	12	Q8B911	Q8b911 porcine rep
24	37	62.7	1463	12	Q9YN01	Q9yn01 porcine rep
25	37	62.7	1463	12	Q9ENK5	Q9enk5 porcine rep
26	37	62.7	1463	12	Q9E8M9	Q9e8m9 porcine rep
27	37	62.7	1463	12	Q99AV5	Q99av5 porcine rep
28	37	62.7	1463	12	Q99BU5	Q99bu5 porcine rep
29	37	62.7	1463	12	Q7TF56	Q7tf56 porcine rep
30	37	62.7	3956	12	Q9DLN9	Q9dln9 porcine rep
31	37	62.7	3960	12	Q9DLP1	Q9dlp1 porcine rep
32	37	62.7	3960	12	Q9DLN8	Q9dln8 porcine rep
33	37	62.7	3960	12	Q9DLP0	Q9dlp0 porcine rep
34	36	61.0	186	16	Q9RUZ7	Q9ruz7 deinococcus
35	36	61.0	244	5	Q8IBP0	Q8ibp0 plasmodium
36	36	61.0	258	6	Q9BEZ5	Q9bez5 choloepus d
37	36	61.0	258	6	Q9BEZ6	Q9bez6 choloepus h
38	36	61.0	258	6	Q9BEZ4	Q9bez4 euphractus
39	36	61.0	457	16	Q89H49	Q89h49 bradyrhizob
40	36	61.0	538	5	Q8IQS7	Q8iqs7 drosophila
41	36	61.0	727	17	Q8PRY3	Q8pry3 methanosarc
42	36	61.0	733	17	Q8TLX6	Q8tlx6 methanosarc
43	36	61.0	4971	5	Q8IBG1	Q8ibg1 plasmodium
44	35	59.3	40	10	Q41267	Q41267 glycine max
45	35	59.3	102	16	Q9JUB1	Q9jub1 neisseria m
46	35	59.3	232	5	Q8ISI2	Q8isi2 oxytricha l
47	35	59.3	278	5	Q9W251	Q9w251 drosophila
48	35	59.3	326	10	Q8W4W4	Q8w4w4 vigna ungui
49	35	59.3	326	10	Q7XZQ1	Q7xzb1 glycine max
50	35	59.3	457	16	Q8PKN1	Q8pkn1 xanthomonas
51	35	59.3	460	10	Q94DF6	Q94df6 oryza sativ
52	35	59.3	730	2	Q9X571	Q9x571 rhizobium l
53	35	59.3	783	16	Q8UBI1	Q8ubi1 agrobacteri
54	35	59.3	857	3	Q8X0W9	Q8x0w9 neurospora
55	35	59.3	927	2	Q87240	Q87240 lactococcus
56	35	59.3	952	5	Q964F9	Q964f9 spodoptera
57	35	59.3	956	13	Q8QG47	Q8qg47 aegithalos

58	35	59.3	1464	3	Q9C229	Q9c229 neurospora
59	34	57.6	101	1	Q8NKU9	Q8nku9 acidianus a
60	34	57.6	102	16	Q9JZ87	Q9jz87 neisseria m
61	34	57.6	106	17	Q97Z50	Q97z50 sulfolobus
62	34	57.6	134	9	Q9G0G7	Q9g0g7 roseophage
63	34	57.6	146	17	Q976Q1	Q976q1 sulfolobus
64	34	57.6	175	10	Q84RF8	Q84rf8 gossypium b
65	34	57.6	223	10	Q9LMB5	Q9lmb5 arabidopsis
66	34	57.6	223	16	Q822Z7	Q822z7 chlamydophi
67	34	57.6	277	16	Q8NRU4	Q8nru4 corynebacte
68	34	57.6	308	2	Q53703	Q53703 staphylococ
69	34	57.6	308	2	Q53731	Q53731 staphylococ
70	34	57.6	308	9	Q9MBN3	Q9mbn3 staphylococ
71	34	57.6	311	2	Q93UU9	Q93uu9 staphylococ
72	34	57.6	311	16	Q99T53	Q99t53 staphylococ
73	34	57.6	314	2	O54081	O54081 staphylococ
74	34	57.6	337	16	Q8CZC1	Q8czc1 streptococc
75	34	57.6	378	12	Q91FM2	Q91fm2 chilo iride
76	34	57.6	379	16	Q8DM04	Q8dm04 synechococc
77	34	57.6	387	16	Q82XE5	Q82xe5 nitrosomona
78	34	57.6	394	6	Q8HZW7	Q8hzw7 ornithorhyn
79	34	57.6	399	16	Q7WR56	Q7wr56 bordetella
80	34	57.6	399	16	Q7W289	Q7w289 bordetella
81	34	57.6	399	16	Q7VT67	Q7vt67 bordetella
82	34	57.6	417	16	Q9KFK7	Q9kfk7 bacillus ha
83	34	57.6	428	16	Q89KR2	Q89kr2 bradyrhizob
84	34	57.6	481	5	Q961I8	Q961i8 drosophila
85	34	57.6	487	16	Q82XZ5	Q82xz5 nitrosomona
86	34	57.6	505	16	Q81SQ9	Q81sq9 bacillus an
87	34	57.6	505	16	Q81FL5	Q81fl5 bacillus ce
88	34	57.6	533	2	Q8GDK2	Q8gdk2 photorhabdu
89	34	57.6	546	16	Q81FG1	Q81fg1 bacillus ce
90	34	57.6	546	16	Q81SI7	Q81si7 bacillus an
91	34	57.6	595	4	Q9UEU4	Q9ueu4 homo sapien
92	34	57.6	597	16	Q9PPX0	Q9ppx0 ureaplasma
93	34	57.6	682	6	Q7YQK2	Q7yqk2 bos taurus
94	34	57.6	773	13	Q8UV20	Q8uv20 sphoeroides
95	34	57.6	855	13	Q802A5	Q802a5 fugu rubrip
96	34	57.6	865	10	Q7X6K3	Q7x6k3 oryza sativ
97	34	57.6	884	10	O23928	O23928 eleocharis
98	34	57.6	951	13	Q8QFY3	Q8qfy3 vanga curvi
99	34	57.6	952	13	Q8QFY5	Q8qfy5 troglodytes
100	34	57.6	952	13	Q8QFY9	Q8qfy9 thamnophilu
101	34	57.6	953	13	Q8QG14	Q8qg14 nectarinia
102	34	57.6	955	13	Q8QG08	Q8qg08 pardalotus
103	34	57.6	956	13	Q8QFY7	Q8qfy7 toxorhamphu
104	34	57.6	956	13	Q9W6P9	Q9w6p9 tyrannus ty
105	34	57.6	957	13	Q8QFZ5	Q8qfz5 regulus cal
106	34	57.6	957	13	Q8QG23	Q8qg23 lanius ludo
107	34	57.6	957	13	Q8QFZ4	Q8qfz4 rupicola ru
108	34	57.6	957	13	Q8QG38	Q8qg38 cinclus cin
109	34	57.6	957	13	Q8QG12	Q8qg12 oriolus lar
110	34	57.6	957	13	Q8QG37	Q8qg37 cisticola a
111	34	57.6	957	13	Q8QG10	Q8qg10 pachycephal
112	34	57.6	957	13	Q8QG19	Q8qg19 menura nova
113	34	57.6	957	13	Q8QFZ1	Q8qfz1 sturnus vul
114	34	57.6	957	13	Q8QG04	Q8qg04 picathartes

115	34	57.6	957	13	Q8QG46	Q8qg46 aegithina t
116	34	57.6	957	13	Q8QG15	Q8qg15 muscicapa s
117	34	57.6	957	13	Q8QG26	Q8qg26 hirundo pyr
118	34	57.6	957	13	Q8QG28	Q8qg28 furnarius r
119	34	57.6	957	13	Q8QG24	Q8qg24 irena cyano
120	34	57.6	957	13	Q8QG21	Q8qg21 melanochari
121	34	57.6	957	13	Q8QG27	Q8qg27 garrulax mi
122	34	57.6	957	13	Q8QG43	Q8qg43 artamus leu
123	34	57.6	957	13	Q8QFZ0	Q8qfz0 sylvia nana
124	34	57.6	957	13	Q8QG13	Q8qg13 oedistoma i
125	34	57.6	957	13	Q8QG30	Q8qg30 formicarius
126	34	57.6	957	13	Q8QG18	Q8qg18 mimus patag
127	34	57.6	957	13	Q8QG03	Q8qg03 pipra coron
128	34	57.6	957	13	Q8QG11	Q8qg11 orthonyx sp
129	34	57.6	957	13	Q8QFY1	Q8qfy1 zosterops s
130	34	57.6	957	13	Q8QG17	Q8qg17 monarcha ax
131	34	57.6	957	13	Q8QG45	Q8qg45 alauda arve
132	34	57.6	957	13	Q8QG35	Q8qg35 coracina li
133	34	57.6	961	13	Q8QG07	Q8qg07 parula amer
134	34	57.6	961	13	Q8QFY8	Q8qfy8 thraupis cy
135	34	57.6	961	13	Q8QG41	Q8qg41 cardinalis
136	34	57.6	961	13	Q8QG31	Q8qg31 emberiza sc
137	34	57.6	961	13	Q9W6P8	Q9w6p8 passer mont
138	34	57.6	1308	10	Q81903	Q81903 arabidopsis
139	34	57.6	1354	11	Q9EPW8	Q9epw8 mus musculu
140	34	57.6	1480	11	Q80TM9	Q80tm9 mus musculu
141	34	57.6	1504	4	Q9UES6	Q9ues6 homo sapien
142	34	57.6	1504	4	Q7Z2X6	Q7z2x6 homo sapien
143	34	57.6	1528	4	Q9Y2I1	Q9y2i1 homo sapien
144	33.5	56.8	387	2	Q8KJM9	Q8kjm9 rhizobium l
145	33.5	56.8	400	16	Q989K2	Q989k2 rhizobium l
146	33	55.9	56	2	Q8RN62	Q8rn62 streptococc
147	33	55.9	157	11	Q8BSY7	Q8bsy7 mus musculu
148	33	55.9	159	10	Q7Y1P6	Q7y1p6 oryza sativ
149	33	55.9	175	16	P73146	P73146 synechocyst
150	33	55.9	175	16	Q89DU3	Q89du3 bradyrhizob

# ALIGNMENTS

## RESULT 1

Q28194

ID Q28194 PRELIMINARY; PRT; 229 AA.  
AC Q28194;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Thrombospondin-1 (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96331130; PubMed=8698834;

RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,  
 RA Feige J.J.;  
 RT "Opposite regulation of thrombospondin-1 and corticotropin-induced  
 RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic  
 RT hormone in adrenocortical cells.";  
 RL J. Cell. Physiol. 167:164-172(1996).  
 DR EMBL; X89511; CAA61682.1; -.  
 DR PIR; S57957; S57957.  
 DR GO; GO:0005198; F:structural molecule activity; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR InterPro; IPR008985; ConA\_like\_lectin.  
 DR InterPro; IPR003129; TSPN.  
 DR Pfam; PF02210; TSPN; 1.  
 DR SMART; SM00210; TSPN; 1.  
 FT NON\_TER 1 1  
 FT NON\_TER 229 229  
 SQ SEQUENCE 229 AA; 25015 MW; 90D9EBCE4E6B669C CRC64;

Query Match 93.2%; Score 55; DB 6; Length 229;  
 Best Local Similarity 100.0%; Pred. No. 0.009;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 QGVQLQNVRFVF 12  
 |||||  
 Db 191 QGVQLQNVRFVF 201

## RESULT 2

Q7SY84

ID Q7SY84 PRELIMINARY; PRT; 496 AA.  
 AC Q7SY84;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:16:36 ; Search time 6.28571 Seconds  
 (without alignments)  
 32.853 Million cell updates/sec

Title: US-10-030-735-51  
 Perfect score: 21  
 Sequence: 1 NVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score	Match					
1	21	100.0	20	5	PCT-US94-10257A-80	Sequence 80, Appl	
2	21	100.0	20	5	PCT-US94-10257A-81	Sequence 81, Appl	
3	21	100.0	36	1	US-08-039-137-41	Sequence 41, Appl	
4	21	100.0	60	4	US-09-621-976-6043	Sequence 6043, Ap	
5	21	100.0	68	4	US-09-134-001C-4018	Sequence 4018, Ap	
6	21	100.0	68	4	US-09-543-681A-7589	Sequence 7589, Ap	
7	21	100.0	74	4	US-09-333-809-153	Sequence 153, App	
8	21	100.0	76	4	US-09-333-809-114	Sequence 114, App	
9	21	100.0	77	4	US-09-333-809-79	Sequence 79, Appl	
10	21	100.0	78	4	US-09-333-809-154	Sequence 154, App	
11	21	100.0	80	4	US-09-489-039A-11771	Sequence 11771, A	

12	21	100.0	80	4	US-09-333-809-118	Sequence 118, App
13	21	100.0	80	4	US-09-333-809-162	Sequence 162, App
14	21	100.0	82	4	US-09-333-809-124	Sequence 124, App
15	21	100.0	82	4	US-09-333-809-136	Sequence 136, App
16	21	100.0	83	4	US-09-333-809-98	Sequence 98, Appl
17	21	100.0	84	4	US-09-333-809-123	Sequence 123, App
18	21	100.0	84	4	US-09-333-809-171	Sequence 171, App
19	21	100.0	85	4	US-09-333-809-29	Sequence 29, Appl
20	21	100.0	85	4	US-09-333-809-90	Sequence 90, Appl
21	21	100.0	85	4	US-09-333-809-117	Sequence 117, App
22	21	100.0	86	4	US-09-333-809-32	Sequence 32, Appl
23	21	100.0	86	4	US-09-333-809-35	Sequence 35, Appl
24	21	100.0	86	4	US-09-333-809-159	Sequence 159, App
25	21	100.0	87	4	US-09-333-809-26	Sequence 26, Appl
26	21	100.0	87	4	US-09-333-809-150	Sequence 150, App
27	21	100.0	88	4	US-09-333-809-128	Sequence 128, App
28	21	100.0	88	4	US-09-333-809-147	Sequence 147, App
29	21	100.0	88	4	US-09-333-809-158	Sequence 158, App
30	21	100.0	89	4	US-09-252-991A-25128	Sequence 25128, A
31	21	100.0	89	4	US-09-333-809-24	Sequence 24, Appl
32	21	100.0	89	4	US-09-333-809-31	Sequence 31, Appl
33	21	100.0	89	4	US-09-333-809-37	Sequence 37, Appl
34	21	100.0	89	4	US-09-333-809-86	Sequence 86, Appl
35	21	100.0	89	4	US-09-333-809-113	Sequence 113, App
36	21	100.0	89	4	US-09-333-809-126	Sequence 126, App
37	21	100.0	89	4	US-09-333-809-134	Sequence 134, App
38	21	100.0	89	4	US-09-333-809-135	Sequence 135, App
39	21	100.0	89	4	US-09-333-809-137	Sequence 137, App
40	21	100.0	89	4	US-09-333-809-139	Sequence 139, App
41	21	100.0	89	4	US-09-333-809-140	Sequence 140, App
42	21	100.0	89	4	US-09-333-809-168	Sequence 168, App
43	21	100.0	89	4	US-09-333-809-170	Sequence 170, App
44	21	100.0	90	4	US-09-333-809-160	Sequence 160, App
45	21	100.0	94	3	US-09-147-550-14	Sequence 14, Appl
46	21	100.0	94	3	US-09-147-550-26	Sequence 26, Appl
47	21	100.0	94	3	US-09-147-550-28	Sequence 28, Appl
48	21	100.0	94	3	US-09-147-550-29	Sequence 29, Appl
49	21	100.0	94	3	US-09-147-550-33	Sequence 33, Appl
50	21	100.0	94	3	US-09-147-550-35	Sequence 35, Appl
51	21	100.0	94	3	US-09-147-550-40	Sequence 40, Appl
52	21	100.0	94	3	US-09-147-550-42	Sequence 42, Appl
53	21	100.0	94	3	US-09-147-550-45	Sequence 45, Appl
54	21	100.0	94	3	US-09-147-550-46	Sequence 46, Appl
55	21	100.0	94	3	US-09-147-550-48	Sequence 48, Appl
56	21	100.0	94	3	US-09-147-550-53	Sequence 53, Appl
57	21	100.0	94	3	US-09-147-550-66	Sequence 66, Appl
58	21	100.0	94	3	US-09-147-550-69	Sequence 69, Appl
59	21	100.0	94	3	US-09-147-550-77	Sequence 77, Appl
60	21	100.0	94	3	US-09-147-550-78	Sequence 78, Appl
61	21	100.0	94	3	US-09-147-550-84	Sequence 84, Appl
62	21	100.0	94	3	US-09-147-550-90	Sequence 90, Appl
63	21	100.0	94	3	US-09-147-550-93	Sequence 93, Appl
64	21	100.0	94	3	US-09-147-550-99	Sequence 99, Appl
65	21	100.0	94	3	US-09-147-550-101	Sequence 101, App
66	21	100.0	94	3	US-09-557-917-14	Sequence 14, Appl
67	21	100.0	94	3	US-09-557-917-26	Sequence 26, Appl
68	21	100.0	94	3	US-09-557-917-28	Sequence 28, Appl



69	21	100.0	94	3	US-09-557-917-29	Sequence 29, Appl
70	21	100.0	94	3	US-09-557-917-33	Sequence 33, Appl
71	21	100.0	94	3	US-09-557-917-35	Sequence 35, Appl
72	21	100.0	94	3	US-09-557-917-40	Sequence 40, Appl
73	21	100.0	94	3	US-09-557-917-42	Sequence 42, Appl
74	21	100.0	94	3	US-09-557-917-45	Sequence 45, Appl
75	21	100.0	94	3	US-09-557-917-46	Sequence 46, Appl
76	21	100.0	94	3	US-09-557-917-48	Sequence 48, Appl
77	21	100.0	94	3	US-09-557-917-53	Sequence 53, Appl
78	21	100.0	94	3	US-09-557-917-66	Sequence 66, Appl
79	21	100.0	94	3	US-09-557-917-69	Sequence 69, Appl
80	21	100.0	94	3	US-09-557-917-77	Sequence 77, Appl
81	21	100.0	94	3	US-09-557-917-78	Sequence 78, Appl
82	21	100.0	94	3	US-09-557-917-84	Sequence 84, Appl
83	21	100.0	94	3	US-09-557-917-90	Sequence 90, Appl
84	21	100.0	94	3	US-09-557-917-93	Sequence 93, Appl
85	21	100.0	94	3	US-09-557-917-99	Sequence 99, Appl
86	21	100.0	94	3	US-09-557-917-101	Sequence 101, App
87	21	100.0	94	4	US-09-333-809-143	Sequence 143, App
88	21	100.0	94	4	US-09-333-809-144	Sequence 144, App
89	21	100.0	96	4	US-09-333-809-165	Sequence 165, App
90	21	100.0	100	1	US-08-241-853-28	Sequence 28, Appl
91	21	100.0	100	1	US-08-241-853-29	Sequence 29, Appl
92	21	100.0	100	2	US-08-850-917-28	Sequence 28, Appl
93	21	100.0	100	2	US-08-850-917-29	Sequence 29, Appl
94	21	100.0	100	4	US-09-333-809-28	Sequence 28, Appl
95	21	100.0	102	4	US-09-333-809-25	Sequence 25, Appl
96	21	100.0	102	4	US-09-333-809-27	Sequence 27, Appl
97	21	100.0	102	4	US-09-333-809-33	Sequence 33, Appl
98	21	100.0	102	4	US-09-333-809-34	Sequence 34, Appl
99	21	100.0	102	4	US-09-333-809-36	Sequence 36, Appl
100	21	100.0	102	4	US-09-333-809-108	Sequence 108, App
101	21	100.0	102	4	US-09-333-809-109	Sequence 109, App
102	21	100.0	102	4	US-09-333-809-197	Sequence 197, App
103	21	100.0	114	4	US-08-747-562-32	Sequence 32, Appl
104	21	100.0	120	2	US-08-535-276-8	Sequence 8, Appli
105	21	100.0	120	2	US-08-535-276-9	Sequence 9, Appli
106	21	100.0	120	4	US-09-335-234-8	Sequence 8, Appli
107	21	100.0	120	4	US-09-335-234-9	Sequence 9, Appli
108	21	100.0	128	4	US-09-134-000C-6625	Sequence 6625, Ap
109	21	100.0	129	1	US-07-820-154A-5	Sequence 5, Appli
110	21	100.0	129	2	US-08-097-554A-5	Sequence 5, Appli
111	21	100.0	129	3	US-08-480-640A-5	Sequence 5, Appli
112	21	100.0	129	3	US-08-295-802-5	Sequence 5, Appli
113	21	100.0	129	3	US-08-686-968C-101	Sequence 101, App
114	21	100.0	129	3	US-08-488-237A-5	Sequence 5, Appli
115	21	100.0	129	4	US-08-375-992A-5	Sequence 5, Appli
116	21	100.0	129	4	US-08-472-679H-5	Sequence 5, Appli
117	21	100.0	129	5	PCT-US93-00324-5	Sequence 5, Appli
118	21	100.0	131	3	US-08-772-440-23	Sequence 23, Appl
119	21	100.0	131	3	US-08-772-440-27	Sequence 27, Appl
120	21	100.0	134	3	US-08-772-440-16	Sequence 16, Appl
121	21	100.0	134	4	US-09-489-039A-8594	Sequence 8594, Ap
122	21	100.0	145	3	US-08-772-440-14	Sequence 14, Appl
123	21	100.0	160	4	US-09-879-833-2	Sequence 2, Appli
124	21	100.0	161	4	US-09-107-532A-5537	Sequence 5537, Ap
125	21	100.0	161	4	US-09-489-039A-9832	Sequence 9832, Ap

126	21	100.0	166	3	US-08-679-006-31	Sequence 31, Appl
127	21	100.0	167	3	US-08-772-440-21	Sequence 21, Appl
128	21	100.0	172	4	US-09-107-532A-4310	Sequence 4310, Ap
129	21	100.0	175	3	US-08-772-440-15	Sequence 15, Appl
130	21	100.0	175	3	US-09-230-637-24	Sequence 24, Appl
131	21	100.0	207	4	US-09-328-352-6694	Sequence 6694, Ap
132	21	100.0	209	3	US-08-772-440-4	Sequence 4, Appli
133	21	100.0	210	3	US-08-737-226-7	Sequence 7, Appli
134	21	100.0	210	4	US-09-254-504-7	Sequence 7, Appli
135	21	100.0	213	4	US-09-489-039A-14310	Sequence 14310, A
136	21	100.0	215	4	US-09-252-991A-26604	Sequence 26604, A
137	21	100.0	218	4	US-10-026-045-1	Sequence 1, Appli
138	21	100.0	218	4	US-10-026-045-2	Sequence 2, Appli
139	21	100.0	218	4	US-10-026-045-3	Sequence 3, Appli
140	21	100.0	220	4	US-09-134-000C-3832	Sequence 3832, Ap
141	21	100.0	221	4	US-09-328-352-5680	Sequence 5680, Ap
142	21	100.0	227	4	US-09-252-991A-27681	Sequence 27681, A
143	21	100.0	228	3	US-09-436-983-7	Sequence 7, Appli
144	21	100.0	236	4	US-09-134-000C-3636	Sequence 3636, Ap
145	21	100.0	242	4	US-09-107-532A-3774	Sequence 3774, Ap
146	21	100.0	247	5	PCT-US94-10257A-2	Sequence 2, Appli
147	21	100.0	250	2	US-08-861-269-5	Sequence 5, Appli
148	21	100.0	250	2	US-09-134-596-5	Sequence 5, Appli
149	21	100.0	250	3	US-09-293-273-5	Sequence 5, Appli
150	21	100.0	259	4	US-09-252-991A-28573	Sequence 28573, A

#### ALIGNMENTS

#### RESULT 1

PCT-US94-10257A-80

; Sequence 80, Application PC/TUS9410257A

; GENERAL INFORMATION:

; APPLICANT: IMMULOGIC PHARMACEUTICAL CORPORATION

; APPLICANT: BRIGITTE DEVAUX

; APPLICANT: JONATHAN B. ROTHBARD

; APPLICANT: DAWN SMILEK

; TITLE OF INVENTION: USES OF MYELIN OLIGODENDROCYTE

; TITLE OF INVENTION: GLYCOPROTEIN AND PEPTIDE PORTIONS THEREOF IN PROTOCOLS

; TITLE OF INVENTION: RELATED TO AUTOIMMUNE DISEASE

; NUMBER OF SEQUENCES: 95

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION

; STREET: 610 LINCOLN STREET

; CITY: WALTHAM

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02145

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: ASCII text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/10257A

; FILING DATE: 1 SEPTEMBER 1994

```

; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/116,824
; FILING DATE: 03-SEP-1993
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: ANNE I CRAIG
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: 071.1 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 466-6000
; TELEFAX: (617) 466-6040
; INFORMATION FOR SEQ ID NO: 80:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 20 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
PCT-US94-10257A-80

```

```

Query Match          100.0%; Score 21; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 39;
Matches      4; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

```

```

Qy      1 NVRF 4
        ||||
Db      17 NVRF 20

```

## RESULT 2

```

PCT-US94-10257A-81
; Sequence 81, Application PC/TUS9410257A
; GENERAL INFORMATION:
; APPLICANT: IMMULOGIC PHARMACEUTICAL CORPORATION
; APPLICANT: BRIGITTE DEVAUX
; APPLICANT: JONATHAN B. ROTHBARD
; APPLICANT: DAWN SMILEK
; TITLE OF INVENTION: USES OF MYELIN OLIGODENDROCYTE
; TITLE OF INVENTION: GLYCOPROTEIN AND PEPTIDE PORTIONS THEREOF IN PROTOCOLS
; TITLE OF INVENTION: RELATED TO AUTOIMMUNE DISEASE
; NUMBER OF SEQUENCES: 95
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: IMMULOGIC PHARMACEUTICAL CORPORATION
; STREET: 610 LINCOLN STREET
; CITY: WALTHAM
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02145
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/10257A
; FILING DATE: 1 SEPTEMBER 1994

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 22:05:26 ; Search time 16.8403 Seconds  
 (without alignments)  
 62.382 Million cell updates/sec

Title: US-10-030-735-51  
 Perfect score: 21  
 Sequence: 1 NVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query					
No.	Score	Match	Length	DB	ID	Description

1	21	100.0	9	12	US-10-253-286-886	Sequence 886, App
2	21	100.0	9	15	US-10-245-871-886	Sequence 886, App
3	21	100.0	10	10	US-09-573-822C-151	Sequence 151, App
4	21	100.0	17	14	US-10-225-567A-1015	Sequence 1015, Ap
5	21	100.0	37	12	US-10-424-599-223109	Sequence 223109,
6	21	100.0	42	12	US-10-424-599-261743	Sequence 261743,
7	21	100.0	45	12	US-10-424-599-227937	Sequence 227937,
8	21	100.0	47	12	US-10-424-599-217193	Sequence 217193,
9	21	100.0	52	12	US-10-424-599-268809	Sequence 268809,
10	21	100.0	54	14	US-10-153-398-7	Sequence 7, Appli
11	21	100.0	56	12	US-10-424-599-157873	Sequence 157873,
12	21	100.0	56	12	US-10-424-599-164123	Sequence 164123,
13	21	100.0	56	12	US-10-424-599-274657	Sequence 274657,
14	21	100.0	57	12	US-10-424-599-181498	Sequence 181498,
15	21	100.0	58	12	US-10-424-599-248538	Sequence 248538,
16	21	100.0	61	12	US-10-424-599-175583	Sequence 175583,
17	21	100.0	62	10	US-09-948-820-70	Sequence 70, Appl
18	21	100.0	63	15	US-10-264-049-4154	Sequence 4154, Ap
19	21	100.0	67	9	US-09-764-869-1088	Sequence 1088, Ap
20	21	100.0	67	14	US-10-091-504-1088	Sequence 1088, Ap
21	21	100.0	67	15	US-10-227-577-1088	Sequence 1088, Ap
22	21	100.0	69	12	US-10-424-599-234228	Sequence 234228,
23	21	100.0	69	12	US-10-424-599-281959	Sequence 281959,
24	21	100.0	69	14	US-10-029-386-29552	Sequence 29552, A
25	21	100.0	70	12	US-10-424-599-205031	Sequence 205031,
26	21	100.0	73	14	US-10-029-386-28872	Sequence 28872, A
27	21	100.0	74	9	US-09-864-761-47701	Sequence 47701, A
28	21	100.0	74	12	US-10-424-599-196587	Sequence 196587,
29	21	100.0	75	12	US-10-424-599-283340	Sequence 283340,
30	21	100.0	75	14	US-10-029-386-33292	Sequence 33292, A
31	21	100.0	83	12	US-10-424-599-170196	Sequence 170196,
32	21	100.0	83	12	US-10-424-599-222704	Sequence 222704,
33	21	100.0	84	12	US-10-424-599-280578	Sequence 280578,
34	21	100.0	89	14	US-10-156-761-8520	Sequence 8520, Ap
35	21	100.0	93	9	US-09-764-869-692	Sequence 692, App
36	21	100.0	93	12	US-10-424-599-188463	Sequence 188463,
37	21	100.0	93	12	US-10-424-599-278205	Sequence 278205,
38	21	100.0	93	14	US-10-091-504-692	Sequence 692, App
39	21	100.0	93	14	US-10-029-386-29273	Sequence 29273, A
40	21	100.0	93	15	US-10-227-577-692	Sequence 692, App
41	21	100.0	95	12	US-10-424-599-243348	Sequence 243348,
42	21	100.0	96	12	US-10-424-599-238659	Sequence 238659,
43	21	100.0	97	12	US-10-424-599-188655	Sequence 188655,
44	21	100.0	97	12	US-10-424-599-220498	Sequence 220498,
45	21	100.0	99	12	US-10-424-599-282132	Sequence 282132,
46	21	100.0	102	15	US-10-630-590-93	Sequence 93, Appl
47	21	100.0	103	9	US-09-764-864-964	Sequence 964, App
48	21	100.0	104	12	US-10-424-599-278081	Sequence 278081,
49	21	100.0	105	12	US-10-282-122A-65700	Sequence 65700, A
50	21	100.0	106	12	US-10-424-599-212728	Sequence 212728,
51	21	100.0	108	14	US-10-083-357-761	Sequence 761, App
52	21	100.0	109	12	US-10-042-865-73	Sequence 73, Appl
53	21	100.0	110	9	US-09-925-297-556	Sequence 556, App
54	21	100.0	110	9	US-09-764-864-1400	Sequence 1400, Ap
55	21	100.0	110	12	US-10-424-599-285315	Sequence 285315,
56	21	100.0	111	12	US-10-425-114-54287	Sequence 54287, A

57	21	100.0	112	10	US-09-764-891-3333	Sequence 3333, Ap
58	21	100.0	112	14	US-10-205-428-372	Sequence 372, App
59	21	100.0	114	15	US-10-349-977-32	Sequence 32, Appl
60	21	100.0	115	14	US-10-029-386-32704	Sequence 32704, A
61	21	100.0	117	9	US-09-738-626-5175	Sequence 5175, Ap
62	21	100.0	117	12	US-10-282-122A-59400	Sequence 59400, A
63	21	100.0	117	12	US-10-282-122A-73214	Sequence 73214, A
64	21	100.0	123	12	US-10-282-122A-75260	Sequence 75260, A
65	21	100.0	123	12	US-10-282-122A-76038	Sequence 76038, A
66	21	100.0	123	12	US-10-424-599-214551	Sequence 214551,
67	21	100.0	123	12	US-10-424-599-259466	Sequence 259466,
68	21	100.0	123	12	US-10-425-114-47956	Sequence 47956, A
69	21	100.0	128	12	US-10-425-114-53883	Sequence 53883, A
70	21	100.0	129	14	US-10-238-075-393	Sequence 393, App
71	21	100.0	132	15	US-10-108-260A-3072	Sequence 3072, Ap
72	21	100.0	133	11	US-09-864-408A-7776	Sequence 7776, Ap
73	21	100.0	133	12	US-10-424-599-265537	Sequence 265537,
74	21	100.0	135	12	US-10-335-977-9190	Sequence 9190, Ap
75	21	100.0	136	12	US-10-335-977-6108	Sequence 6108, Ap
76	21	100.0	136	12	US-10-335-977-9191	Sequence 9191, Ap
77	21	100.0	142	14	US-10-080-170-210	Sequence 210, App
78	21	100.0	142	14	US-10-080-170-416	Sequence 416, App
79	21	100.0	150	15	US-10-369-493-13190	Sequence 13190, A
80	21	100.0	158	12	US-10-424-599-185902	Sequence 185902,
81	21	100.0	158	12	US-10-424-599-234976	Sequence 234976,
82	21	100.0	160	9	US-09-879-833-2	Sequence 2, Appli
83	21	100.0	160	9	US-09-779-307-27	Sequence 27, Appl
84	21	100.0	160	12	US-10-114-270-88	Sequence 88, Appl
85	21	100.0	161	10	US-09-766-511B-55	Sequence 55, Appl
86	21	100.0	163	12	US-10-042-865-75	Sequence 75, Appl
87	21	100.0	169	12	US-10-424-599-175915	Sequence 175915,
88	21	100.0	170	12	US-10-463-957-7	Sequence 7, Appli
89	21	100.0	175	12	US-10-003-632C-5	Sequence 5, Appli
90	21	100.0	181	15	US-10-108-260A-4660	Sequence 4660, Ap
91	21	100.0	182	12	US-10-424-599-180818	Sequence 180818,
92	21	100.0	183	12	US-10-042-865-74	Sequence 74, Appl
93	21	100.0	185	9	US-09-738-626-6392	Sequence 6392, Ap
94	21	100.0	186	12	US-10-282-122A-65012	Sequence 65012, A
95	21	100.0	191	11	US-09-801-944B-147	Sequence 147, App
96	21	100.0	193	12	US-10-424-599-232182	Sequence 232182,
97	21	100.0	194	9	US-09-734-017A-74	Sequence 74, Appl
98	21	100.0	197	12	US-10-282-122A-51966	Sequence 51966, A
99	21	100.0	198	10	US-09-966-422B-12	Sequence 12, Appl
100	21	100.0	198	12	US-10-424-599-199854	Sequence 199854,
101	21	100.0	198	14	US-10-262-272A-12	Sequence 12, Appl
102	21	100.0	200	12	US-09-925-298-714	Sequence 714, App
103	21	100.0	200	14	US-10-102-806-714	Sequence 714, App
104	21	100.0	202	12	US-10-042-865-8	Sequence 8, Appli
105	21	100.0	202	12	US-10-081-056-368	Sequence 368, App
106	21	100.0	202	14	US-10-223-085-368	Sequence 368, App
107	21	100.0	202	14	US-10-223-084-368	Sequence 368, App
108	21	100.0	202	14	US-10-223-088-368	Sequence 368, App
109	21	100.0	202	14	US-10-223-090-368	Sequence 368, App
110	21	100.0	202	14	US-10-223-087-368	Sequence 368, App
111	21	100.0	202	14	US-10-223-083-368	Sequence 368, App
112	21	100.0	202	14	US-10-223-089-368	Sequence 368, App
113	21	100.0	202	14	US-10-223-081-368	Sequence 368, App

114	21	100.0	202	14	US-10-223-082-368	Sequence 368, App
115	21	100.0	202	15	US-10-291-265-267	Sequence 267, App
116	21	100.0	204	12	US-10-282-122A-44916	Sequence 44916, A
117	21	100.0	204	12	US-10-282-122A-44979	Sequence 44979, A
118	21	100.0	206	12	US-10-424-599-146713	Sequence 146713, A
119	21	100.0	209	10	US-09-766-511B-53	Sequence 53, Appl
120	21	100.0	209	10	US-09-766-511B-60	Sequence 60, Appl
121	21	100.0	209	14	US-10-270-470-4	Sequence 4, Appli
122	21	100.0	209	14	US-10-270-470-10	Sequence 10, Appl
123	21	100.0	210	12	US-10-424-599-174096	Sequence 174096, A
124	21	100.0	210	12	US-10-425-114-37706	Sequence 37706, A
125	21	100.0	210	14	US-10-310-630-7	Sequence 7, Appli
126	21	100.0	211	12	US-10-282-122A-56924	Sequence 56924, A
127	21	100.0	211	12	US-10-424-599-175602	Sequence 175602, A
128	21	100.0	212	12	US-10-425-114-68027	Sequence 68027, A
129	21	100.0	217	12	US-10-282-122A-57089	Sequence 57089, A
130	21	100.0	217	15	US-10-369-493-8033	Sequence 8033, Ap
131	21	100.0	221	12	US-10-424-599-245858	Sequence 245858, A
132	21	100.0	223	12	US-10-425-114-65736	Sequence 65736, A
133	21	100.0	224	10	US-09-882-227-40	Sequence 40, Appl
134	21	100.0	224	12	US-10-282-122A-67198	Sequence 67198, A
135	21	100.0	224	12	US-10-335-977-5652	Sequence 5652, Ap
136	21	100.0	228	12	US-10-424-599-243919	Sequence 243919, A
137	21	100.0	228	12	US-10-335-977-5653	Sequence 5653, Ap
138	21	100.0	228	14	US-10-264-820-7	Sequence 7, Appli
139	21	100.0	228	14	US-10-291-190-18	Sequence 18, Appl
140	21	100.0	228	14	US-10-291-190-22	Sequence 22, Appl
141	21	100.0	232	12	US-10-424-599-191891	Sequence 191891, A
142	21	100.0	232	15	US-10-369-493-2302	Sequence 2302, Ap
143	21	100.0	233	15	US-10-369-493-4081	Sequence 4081, Ap
144	21	100.0	233	15	US-10-369-493-19746	Sequence 19746, A
145	21	100.0	236	12	US-10-425-114-71107	Sequence 71107, A
146	21	100.0	237	12	US-10-424-599-221068	Sequence 221068, A
147	21	100.0	237	15	US-10-369-493-11120	Sequence 11120, A
148	21	100.0	237	15	US-10-369-493-23677	Sequence 23677, A
149	21	100.0	238	12	US-10-425-114-58653	Sequence 58653, A
150	21	100.0	240	12	US-10-419-462-40	Sequence 40, Appl

#### ALIGNMENTS

##### RESULT 1

US-10-253-286-886

; Sequence 886, Application US/10253286

; Publication No. US20040058881A1

; GENERAL INFORMATION:

; APPLICANT: HUMPHREYS, ROBERT

; APPLICANT: XU, MINZHEN

; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES

; FILE REFERENCE: REH-2015

; CURRENT APPLICATION NUMBER: US/10/253,286

; CURRENT FILING DATE: 2003-01-13

; PRIOR APPLICATION NUMBER: 10/197,000

; PRIOR FILING DATE: 2002-07-17

; PRIOR APPLICATION NUMBER: 09/396,813

; PRIOR FILING DATE: 1999-09-14

; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 886  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-253-286-886

Query Match 100.0%; Score 21; DB 12; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVRF 4  
|||  
Db 5 NVRF 8

RESULT 2

US-10-245-871-886  
; Sequence 886, Application US/10245871  
; Publication No. US20030235594A1  
; GENERAL INFORMATION:  
; APPLICANT: HUMPHREYS, ROBERT  
; APPLICANT: XU, MINZHEN  
; TITLE OF INVENTION: Ii-KEY/ANTIGENIC EPITOPE HYBRID PEPTIDE VACCINES  
; FILE REFERENCE: REH-2013  
; CURRENT APPLICATION NUMBER: US/10/245,871  
; CURRENT FILING DATE: 2003-01-09  
; PRIOR APPLICATION NUMBER: 10/197,000  
; PRIOR FILING DATE: 2002-07-17  
; PRIOR APPLICATION NUMBER: 09/396,813  
; PRIOR FILING DATE: 1999-09-14  
; NUMBER OF SEQ ID NOS: 905  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 886  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-245-871-886

Query Match 100.0%; Score 21; DB 15; Length 9;  
Best Local Similarity 100.0%; Pred. No. 9.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVRF 4  
|||  
Db 5 NVRF 8

RESULT 3

US-09-573-822C-151  
; Sequence 151, Application US/09573822C  
; Publication No. US20030199011A1  
; GENERAL INFORMATION:  
; APPLICANT: Proteom Ltd  
; TITLE OF INVENTION: Complementary peptide ligands generated from microbial genome sequences



OM protein - protein search, using sw model

Run on: April 7, 2004, 19:11:45 ; Search time 5.10924 Seconds  
 (without alignments)  
 75.308 Million cell updates/sec

Title: US-10-030-735-51  
 Perfect score: 21  
 Sequence: 1 NVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	No.	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	17	2	A27636	cytotoxin B - Clos	
2	21	100.0	39	2	G97623	hypothetical prote	
3	21	100.0	69	2	I59630	MHC cell surface g	
4	21	100.0	70	2	F64066	probable outer mem	
5	21	100.0	72	2	PH0155	HLA-DRB sigma anti	
6	21	100.0	72	2	PH0158	HLA-DRB sigma anti	
7	21	100.0	72	2	PT0162	HLA-DRB sigma anti	
8	21	100.0	72	2	PT0170	HLA-DRB sigma anti	
9	21	100.0	72	2	PT0171	HLA-DRB sigma anti	
10	21	100.0	72	2	B85608	hypothetical prote	
11	21	100.0	73	2	C33287	MHC class II histo	
12	21	100.0	73	2	I59625	MHC class II HLA-D	
13	21	100.0	73	2	I79659	DRB1 transplantati	

14	21	100.0	73	2	I79660	DRB1 transplantati
15	21	100.0	74	2	I68827	MHC class II HLA-D
16	21	100.0	75	2	AD2183	hypothetical prote
17	21	100.0	78	2	I45929	MHC class II lymph
18	21	100.0	78	2	I54528	MHC class II HLA-D
19	21	100.0	80	2	S54907	Beta 1 domain of M
20	21	100.0	81	2	I47060	MHC class II antig
21	21	100.0	81	2	S38696	class II histocomp
22	21	100.0	82	2	T15935	hypothetical prote
23	21	100.0	83	2	S38698	class II histocomp
24	21	100.0	83	2	S38694	class II histocomp
25	21	100.0	84	2	G01972	cGMP-phosphodieste
26	21	100.0	84	2	AD2824	conserved hypothet
27	21	100.0	85	2	S03439	MHC class II histo
28	21	100.0	85	2	B45832	MHC class II histo
29	21	100.0	85	2	I59647	gene HLA-DRB1 prot
30	21	100.0	86	2	C28515	MHC class II histo
31	21	100.0	86	2	AH2549	hypothetical prote
32	21	100.0	88	2	T15898	hypothetical prote
33	21	100.0	89	2	A33287	MHC class II histo
34	21	100.0	89	2	I68554	cell surface glyco
35	21	100.0	90	2	B27060	class II histocomp
36	21	100.0	90	2	B99799	hypothetical protei
37	21	100.0	90	2	G84514	probable glycine-r
38	21	100.0	93	2	S27028	hypothetical prote
39	21	100.0	93	2	S78237	apodystrophin 1 -
40	21	100.0	94	2	B30575	MHC class II histo
41	21	100.0	96	2	G69745	hypothetical prote
42	21	100.0	96	4	S53949	hypothetical prote
43	21	100.0	98	2	AG2440	hypothetical prote
44	21	100.0	105	2	E81956	PEMK-like protein
45	21	100.0	107	2	C81014	PemK-related prote
46	21	100.0	107	2	T51124	lycopene cyclase [
47	21	100.0	108	2	E69358	conserved hypothet
48	21	100.0	112	2	T41131	very hypothetical
49	21	100.0	112	2	H95001	IS630-Spn1, transp
50	21	100.0	112	2	H95235	IS630-Spn1, transp
51	21	100.0	114	2	H90363	conserved hypothet
52	21	100.0	114	2	A05236	hypothetical prote
53	21	100.0	116	2	A86799	hypothetical prote
54	21	100.0	118	2	D61409	genome polyprotein
55	21	100.0	121	2	C81401	probable flagellar
56	21	100.0	123	2	C25239	MHC class II histo
57	21	100.0	123	2	S26610	transcription fact
58	21	100.0	123	2	AF1045	probable membrane
59	21	100.0	125	2	E81783	hypothetical prote
60	21	100.0	125	2	H81206	hypothetical prote
61	21	100.0	126	2	T20427	hypothetical prote
62	21	100.0	126	2	I60079	opsin, middle-wave
63	21	100.0	126	2	AB3177	conserved hypothet
64	21	100.0	128	2	S76468	hypothetical prote
65	21	100.0	128	2	C97602	yefm protein [impo
66	21	100.0	131	2	T17545	hypothetical prote
67	21	100.0	132	2	G82713	30S ribosomal prot
68	21	100.0	139	2	G90414	conserved hypothet
69	21	100.0	141	2	S45450	DER1 protein - yea
70	21	100.0	142	2	G87147	conserved hypothet

71	21	100.0	142	2	A70613	hypothetical prote
72	21	100.0	144	2	D89934	hypothetical prote
73	21	100.0	146	2	H86903	hypothetical prote
74	21	100.0	148	2	T37532	ubiquitin-conjugat
75	21	100.0	149	2	AF2713	conserved hypothet
76	21	100.0	160	1	S17194	endothelin 2 precu
77	21	100.0	162	2	H83459	cytochrome C-type
78	21	100.0	162	2	D82802	hypothetical prote
79	21	100.0	164	2	E97061	integrase XerD fam
80	21	100.0	164	2	C97495	hypothetical prote
81	21	100.0	167	2	T50036	beta-1,4-galactosy
82	21	100.0	170	1	JC4332	phospholipid-hydro
83	21	100.0	181	2	S74968	hypothetical prote
84	21	100.0	183	2	A37410	H-2 class II histo
85	21	100.0	183	2	I50442	VMO-I - chicken
86	21	100.0	184	2	T23518	heterochromatin pr
87	21	100.0	187	2	T49684	hypothetical prote
88	21	100.0	189	2	G82586	translation elonga
89	21	100.0	190	2	S28605	GTP-binding protei
90	21	100.0	191	2	S29558	site-specific DNA-
91	21	100.0	192	2	S77023	hypothetical prote
92	21	100.0	193	2	A64397	hypothetical prote
93	21	100.0	194	2	JQ1095	hypothetical 21.8K
94	21	100.0	196	2	S07728	hypothetical prote
95	21	100.0	197	2	D97061	uncharacterized lo
96	21	100.0	199	2	S26355	opsin, 556nm - whi
97	21	100.0	199	2	S26356	opsin, 543nm - whi
98	21	100.0	199	2	D36131	sigma-B activity i
99	21	100.0	199	2	S77228	hypothetical prote
100	21	100.0	201	2	S52833	probable heme bind
101	21	100.0	202	2	T50101	adenylylsulfate ki
102	21	100.0	210	2	S32822	repressor protein
103	21	100.0	212	2	T20295	hypothetical prote
104	21	100.0	213	2	C83772	hypothetical prote
105	21	100.0	218	2	B47712	myelin/oligodendro
106	21	100.0	219	2	G95913	probable cell surf
107	21	100.0	219	2	D90448	conserved hypothet
108	21	100.0	219	2	T24429	hypothetical prote
109	21	100.0	220	2	A25925	class II histocomp
110	21	100.0	220	2	I68710	HLA-DR beta-chain
111	21	100.0	224	1	A64599	hydrogenase (EC 1.
112	21	100.0	224	2	D71915	hydrogenase, cytoc
113	21	100.0	224	2	I37243	CMRF-35 antigen -
114	21	100.0	228	2	E69758	hypothetical prote
115	21	100.0	229	2	S57957	thrombospondin 1 -
116	21	100.0	232	1	JH0597	transcription fact
117	21	100.0	233	2	JQ0097	hypothetical 26K p
118	21	100.0	237	2	B91282	probable oxidoredu
119	21	100.0	237	2	D86123	probable oxidoredu
120	21	100.0	237	2	D65237	hypothetical 24.6
121	21	100.0	237	2	C27060	class II histocomp
122	21	100.0	237	2	H28043	MHC class II histo
123	21	100.0	237	2	I54294	MHC HLA-DR-beta su
124	21	100.0	239	2	H90218	SSU ribosomal prot
125	21	100.0	240	2	T41254	probable coatmer d
126	21	100.0	241	2	T20432	hypothetical prote
127	21	100.0	241	2	F81835	hypothetical prote

128	21	100.0	241	2	B81129	conserved hypothet
129	21	100.0	242	2	AC3287	hypothetical cytos
130	21	100.0	242	2	T24034	hypothetical prote
131	21	100.0	243	2	F71861	hypothetical prote
132	21	100.0	245	2	A87329	hypothetical prote
133	21	100.0	246	2	S63639	hypothetical prote
134	21	100.0	246	2	A47712	myelin/oligodendro
135	21	100.0	247	2	A55717	myelin/oligodendro
136	21	100.0	247	2	S58394	myelin/oligodendro
137	21	100.0	248	2	D90897	probable oxidoredu
138	21	100.0	248	2	C85720	probable oxidoredu
139	21	100.0	248	2	F64908	probable dehydroge
140	21	100.0	250	2	S36769	ubiquitin-protein
141	21	100.0	252	2	H64122	ydfG protein - Hae
142	21	100.0	253	2	G81389	probable DNA-direc
143	21	100.0	256	2	S37926	hypothetical prote
144	21	100.0	258	2	S18001	ribosomal protein
145	21	100.0	258	2	A64318	hypothetical prote
146	21	100.0	261	2	C86207	hypothetical prote
147	21	100.0	262	2	A75009	probable aryl phos
148	21	100.0	262	2	A88482	protein C05D11.5 [
149	21	100.0	263	2	A35147	hypothetical prote
150	21	100.0	264	2	E82903	1-acyl-sn-glycerol

#### ALIGNMENTS

##### RESULT 1

A27636

cytotoxin B - Clostridium difficile (fragment)

C;Species: Clostridium difficile

C;Date: 15-Dec-1988 #sequence\_revision 15-Dec-1988 #text\_change 18-Jun-1993

C;Accession: A27636

R;Meador III, J.; Tweten, R.K.

Infect. Immun. 56, 1708-1714, 1988

A;Title: Purification and characterization of toxin B from Clostridium difficile.

A;Reference number: A27636; MUID:88256296; PMID:3384474

A;Accession: A27636

A;Molecule type: protein

A;Residues: 1-17 <MEA>

C;Keywords: cytotoxin

Query Match 100.0%; Score 21; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 17;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVRF 4  
 ||||  
 Db 13 NVRF 16

##### RESULT 2

G97623

hypothetical protein AGR\_C\_4001 [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C;Species: Agrobacterium tumefaciens  
 C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002  
 C;Accession: G97623  
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.;  
 Goldman, B.S.; Cao, Y.; Askenazi, M.; Halling, C.; Mullin, L.; Houmiel, K.;  
 Gordon, J.; Vaudin, M.; Iartchouk, O.; Epp, A.; Liu, F.; Wollam, C.; Allinger,  
 M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Flanagan, C.; Crowell, C.;  
 Gurson, J.; Lomo, C.; Sear, C.; Strub, G.; Cielo, C.; Slater, S.  
 Science 294, 2323-2328, 2001  
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent  
 Agrobacterium tumefaciens C58.  
 A;Reference number: A97359; MUID:21608551; PMID:11743194  
 A;Accession: G97623  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-39 <KUR>  
 A;Cross-references: GB:AE007869; PIDN:AAK87944.1; PID:g15157348; GSPDB:GN00169  
 C;Genetics:  
 A;Gene: AGR\_C\_4001  
 A;Map position: circular chromosome

Query Match 100.0%; Score 21; DB 2; Length 39;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVRF 4  
 ||||  
 Db 21 NVRF 24

# RESULT 3

I59630  
 MHC cell surface glycoprotein - human (fragment)  
 C;Species: Homo sapiens (man)  
 C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 23-Jul-1999  
 C;Accession: I59630  
 R;Apple, R.J.; Erlich, H.A.  
 Tissue Antigens 40, 69-74, 1992  
 A;Title: Two new HLA DRB1 alleles found in African Americans: implications for  
 balancing selection at positions 57 and 86.  
 A;Reference number: I59630; MUID:93031787; PMID:1412419  
 A;Accession: I59630  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-69 <RES>  
 A;Cross-references: GB:M81743; NID:g188279; PIDN:AAA36284.1; PID:g188280  
 C;Superfamily: class II histocompatibility antigen; immunoglobulin homology  
 C;Keywords: glycoprotein

Query Match 100.0%; Score 21; DB 2; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVRF 4  
 ||||  
 Db 28 NVRF 31

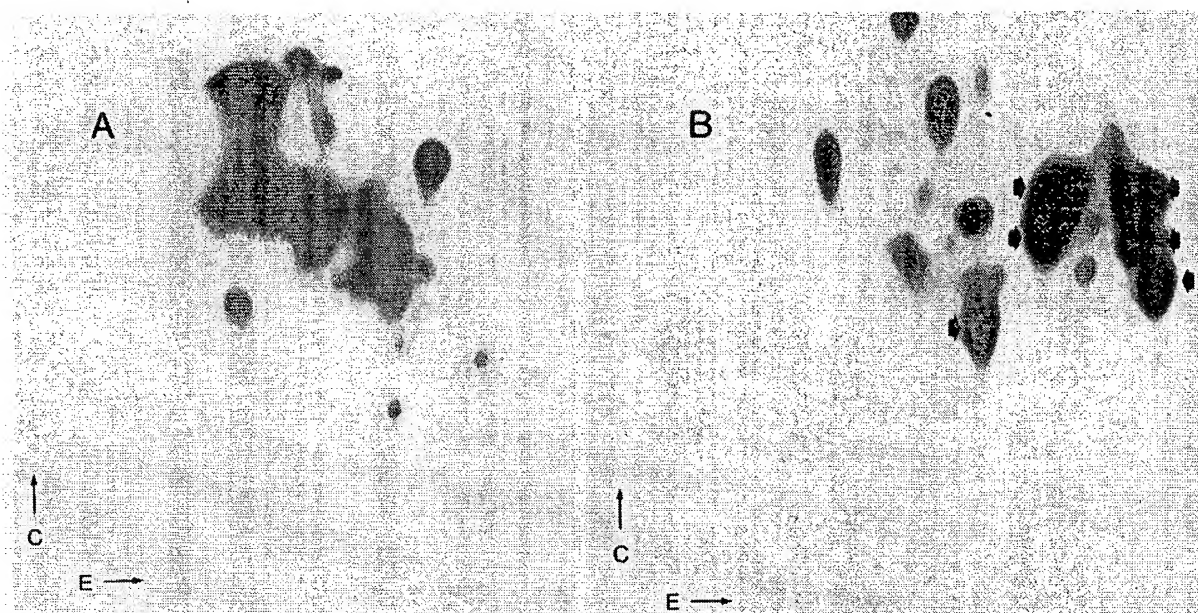


FIG. 8. Analysis of tryptic peptides from toxin B and the 50-kDa protein contaminant. The two-dimensional tryptic maps of the 250-kDa toxin B (A) and the 50-kDa contaminating protein (B) are shown. E and C indicate the directions of the electrophoresis and chromatography steps, respectively, and the heavy arrows in panel B indicate the major peptides which are unique to the 50-kDa protein.

toxin B increased slightly after the 50-kDa protein was removed. Also, the 50-kDa protein was not derived from the 250-kDa toxin by proteolytic cleavage. The fact that the 50-kDa protein copurified with toxin B was due to similarities in the physical characteristics of the native 50-kDa protein and toxin B. A weak interaction between toxin B and the 50-kDa protein, however, cannot be ruled out.

Analyses of toxin B and the 50-kDa protein by SDS-PAGE and native PAGE reveal a significant difference in the mobility of the 50-kDa protein in these two gel methods. Separation of the 50-kDa protein from toxin B was easily achieved by SDS-PAGE under either reducing or nonreducing conditions. In contrast, the migrations of both proteins were nearly identical on native PAGE. This was not surprising, since they appear to be similar in their charge and native molecular weight as judged by their retention characteristics on the high-resolution anion-exchange and gel filtration columns. In view of these observations, native PAGE is not appropriate for the comparison of toxin B preparations, since toxin B and the 50-kDa protein have nearly identical mobilities in this system.

Analysis of purified toxin B by high-resolution gel filtration revealed that the apparent native molecular weight was about 500,000. One possible explanation of this result is that toxin B normally forms dimers in solution. Whether dimerization is actually responsible for the apparent molecular weight of the native toxin B is currently unknown, although this confirms the observation of workers who have reported the molecular weight of toxin B to be between 400,000 and 600,000 (2). Toxin B prepared in this study was highly cytotoxic when applied to sensitive cells. One cytotoxic unit was estimated to be 0.2 to 0.8 pg of purified toxin B. This value is similar to that obtained by Banno et al. (2) and Lysterly et al. (6) for their preparations of toxin B.

The findings of this study help resolve some of the discrepancies in the reported description of toxin B from *C. difficile*. The 50-kDa protein exhibited the properties of the

protein isolated by some investigators (9–11) as toxin B. This protein is a contaminant, but under nondenaturing conditions it has physical properties similar to those of toxin B. We also provide a standard purification scheme for the preparation of highly active, homogeneous toxin B and the amino-terminal sequence of toxin B. The availability of homogeneous toxin B will be important in studies which are aimed at understanding the mechanism and structure-function relationships of toxin B.

#### ACKNOWLEDGMENTS

We thank R. McCallum and J. Ferretti for critical reading of the manuscript.

This work was supported in part by a grant from the Research Council of the University of Oklahoma Health Sciences Center.

#### LITERATURE CITED

1. Aswell, J. E., E. Ehlich, R. Van Tassel, C. C. Tsai, L. V. Holdeman, and T. D. Wilkins. 1979. Characterization and comparison of *Clostridium difficile* and other clostridial toxins, p. 272–275. In D. Schlessinger (ed.), *Microbiology—1979*. American Society for Microbiology, Washington, D.C.
2. Banno, Y., T. Kobayashi, K. Watanabe, K. Ueno, and Y. Nozawa. 1984. Biochemical characterization and biologic actions of two toxins (D1 and D2) from *Clostridium difficile*. *Rev. Infect. Dis.* 6:S11–S20.
3. Bradford, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72:248–254.
4. Krause, R. M., and M. McCarty. 1962. Studies in the chemical structure of the streptococcal cell wall. II. The composition of the group C cell walls and the chemical basis for serological specificity of the carbohydrate moiety. *J. Exp. Med.* 115:49–62.
5. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680–685.
6. Lysterly, D. M., M. D. Roberts, C. J. Phelps, and T. D. Wilkins. 1986. Purification and properties of toxins A and B of *Clostridium difficile*. *FEMS Microbiol. Lett.* 33:31–35.

7. Lyster, D. M., K. E. Saum, D. K. McDonald, and T. D. Wilkins. 1985. Effects of *Clostridium difficile* toxins given intragastrically to animals. *Infect. Immun.* 47:349-352.
8. Mitchell, M. J., B. E. Laughon, and S. Lin. 1987. Biochemical studies on the effect of *Clostridium difficile* toxin B on actin in vivo and in vitro. *Infect. Immun.* 55:1610-1615.
9. Pothoulakis, C., L. M. Barone, R. Ely, B. Faris, M. E. Clark, C. Franzblau, and J. T. LaMont. 1986. Purification and properties of *Clostridium difficile* cytotoxin. *J. Biol. Chem.* 261:1316-1321.
10. Rihn, B., and H. Montell. 1984. Cytotoxin of *Clostridium difficile*, p. 377-378. *In* J. E. Alouf, F. J. Fehrenbach, J. H. Freer, and J. J. Jelszewicz, (ed.), *Bacterial protein toxins*. Academic Press, Inc. (London), Ltd., London.
11. Rolfe, R. D., and S. M. Finegold. 1979. Purification and characterization of *Clostridium difficile* toxin. *Infect. Immun.* 25:191-201.
12. Sullivan, N. M., S. Pellett, and T. D. Wilkins. 1982. Purification and characterization of toxins A and B of *Clostridium difficile*. *Infect. Immun.* 35:1032-1040.
13. Taylor, N. S., and J. G. Bartlett. 1979. Partial purification and characterization of a cytotoxin from *Clostridium difficile*. *Rev. Infect. Dis.* 1:379-385.
14. Taylor, N. S., G. M. Thorne, and J. G. Bartlett. 1981. Comparison of two toxins produced by *Clostridium difficile*. *Infect. Immun.* 34:1036-1043.
15. Tweten, R. K., and J. J. Iandolo. 1981. Purification and partial characterization of a putative precursor to staphylococcal enterotoxin B. *Infect. Immun.* 34:900-907.

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:58:36 ; Search time 2.72269 Seconds  
 (without alignments)  
 76.498 Million cell updates/sec

Title: US-10-030-735-51  
 Perfect score: 21  
 Sequence: 1 NVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	21	100.0	20	1	RIPX_CUCPE	P80750 cucurbita p
2	21	100.0	70	1	Y414_HAEIN	Q57392 haemophilus
3	21	100.0	93	1	YCX1_PAVLU	Q01572 pavlova lut
4	21	100.0	94	1	RPC1_BP434	P16117 bacterioph
5	21	100.0	96	1	YBCH_BACSU	O34795 bacillus su
6	21	100.0	106	1	YX83_METAC	Q8tkm0 methanosarc
7	21	100.0	108	1	Y869_ARCFU	O29392 archaeoglob
8	21	100.0	114	1	YJ76_SULSO	Q97wy4 sulfolobus
9	21	100.0	119	1	Y182_THEAC	Q9hlp5 thermoplasm
10	21	100.0	132	1	RS8_XYLFA	Q9pe62 xylella fas
11	21	100.0	132	1	RS8_XYLFT	Q87e68 xylella fas
12	21	100.0	148	1	UBCC_SCHPO	O13685 schizosacch
13	21	100.0	159	1	RODL_ASPFU	P41746 aspergillus
14	21	100.0	174	1	FLAV_BUCBP	Q89ak0 buchnera ap
15	21	100.0	183	1	VMO1_CHICK	P41366 gallus gall
16	21	100.0	188	1	EFPL_XANAC	Q8plf1 xanthomonas
17	21	100.0	188	1	EFPL_XANCP	Q8p9m3 xanthomonas



18	21	100.0	188	1	EFPL_XYLFT	Q87c43	xylella fas
19	21	100.0	189	1	EFPL_XYLFA	Q9pbe1	xylella fas
20	21	100.0	189	1	SAR1_TRIRE	P78976	trichoderma
21	21	100.0	190	1	SAR1_PICPA	Q9p4c8	pichia past
22	21	100.0	190	1	SAR1_SCHPO	Q01475	schizosacch
23	21	100.0	191	1	Y777_METJA	Q58187	methanococc
24	21	100.0	194	1	YR22_TRSVR	P25246	tomato ring
25	21	100.0	196	1	YM02_PARTE	P15604	paramecium
26	21	100.0	197	1	GX41_RAT	P36970	rattus norv
27	21	100.0	199	1	RSBX_BACSU	P17906	bacillus su
28	21	100.0	201	1	YMW3_YEAST	Q04772	saccharomyc
29	21	100.0	202	1	KAPS_SCHPO	Q9p7g9	schizosacch
30	21	100.0	206	1	KAPS_EMENI	Q92203	emer icella
31	21	100.0	211	1	DER1_YEAST	P38307	saccharomyc
32	21	100.0	212	1	NOS_SQUAC	Q9i9m2	squalus aca
33	21	100.0	214	1	CAPA_KLULA	O74232	kluyveromyc
34	21	100.0	218	1	HIS5_HALN1	Q9hni6	halobacteri
35	21	100.0	224	1	CM35_HUMAN	Q08708	homo sapien
36	21	100.0	224	1	RADC_PASMU	P57913	pasteurella
37	21	100.0	233	1	VHEL_PMV	P20952	papaya mosa
38	21	100.0	237	1	YJGI_ECOLI	P39333	escherichia
39	21	100.0	239	1	RS4E_SULSO	Q9ux94	sulfolobus
40	21	100.0	245	1	MOG_RAT	Q63345	rattus norv
41	21	100.0	246	1	MOG_BOVIN	P55803	bos taurus
42	21	100.0	246	1	MOG_MOUSE	Q61885	mus musculu
43	21	100.0	247	1	MOG_HUMAN	Q16653	homo sapien
44	21	100.0	248	1	YDFG_ECOLI	P39831	escherichia
45	21	100.0	250	1	UBC6_YEAST	P33296	saccharomyc
46	21	100.0	252	1	YDFG_HAEIN	P45200	haemophilus
47	21	100.0	253	1	GX42_RAT	Q91xr8	rattus norv
48	21	100.0	256	1	UT11_YEAST	P34247	saccharomyc
49	21	100.0	258	1	RK15_PEA	P31165	pisum sativ
50	21	100.0	258	1	Y144_METJA	Q57608	methanococc
51	21	100.0	261	1	HIS6_DEIRA	Q9rwd7	deinococcus
52	21	100.0	262	1	YPD5_CAEEL	Q11185	caenorhabdi
53	21	100.0	263	1	YXIS_SACER	P22875	saccharopol
54	21	100.0	265	1	GCTB_ACIFE	Q59112	acidaminoco
55	21	100.0	266	1	2B19_HUMAN	Q9tqe0	homo sapien
56	21	100.0	266	1	HB2B_HUMAN	P01912	homo sapien
57	21	100.0	272	1	Y603_AQUAE	O66861	aquifex aeo
58	21	100.0	273	1	OPSG_ODOVI	O18911	odocoileus
59	21	100.0	273	1	OPSR_CANFA	O18914	canis famil
60	21	100.0	273	1	OPSR_HORSE	O18912	equus cabal
61	21	100.0	280	1	MNTC_LISIN	Q92ag0	listeria in
62	21	100.0	287	1	NARA_MOUSE	P17981	mus musculu
63	21	100.0	288	1	HEMK_BACSU	P45873	bacillus su
64	21	100.0	291	1	CU59_DROME	Q9vzh1	drosophila
65	21	100.0	299	1	HEM6_ECO57	Q8xbi4	escherichia
66	21	100.0	299	1	HEM6_ECOL6	Q8ffa3	escherichia
67	21	100.0	299	1	HEM6_ECOLI	P36553	escherichia
68	21	100.0	299	1	HEM6_SALTI	Q8z4u8	salmonella
69	21	100.0	299	1	HEM6_SALTY	P33771	salmonella
70	21	100.0	303	1	HEM6_PSEPK	Q88rq6	pseudomonas
71	21	100.0	304	1	HEM6_PSESM	Q88b49	pseudomonas
72	21	100.0	305	1	HEM6_PSEAE	P43898	pseudomonas
73	21	100.0	305	1	HEM6_VIBCH	Q9kvt4	vibrio chol
74	21	100.0	305	1	HEM6_VIBPA	Q87ke3	vibrio para

75	21	100.0	305	1	HEM6_VIBVU	Q8ddd5	vibrio vuln
76	21	100.0	309	1	HEM6_YERPE	Q8zcf9	yersinia pe
77	21	100.0	311	1	PUR2_VIGUN	P52421	vigna ungui
78	21	100.0	323	1	YK77_MYCTU	Q10685	mycobacteri
79	21	100.0	325	1	YJB7_YEAST	P47071	saccharomyc
80	21	100.0	326	1	AKD1_HUMAN	P51857	homo sapien
81	21	100.0	326	1	AKD1_RAT	P31210	rattus norv
82	21	100.0	329	1	YX99_BACHD	Q9k7g2	bacillus ha
83	21	100.0	334	1	ARGC_BUCAI	P57156	buchnera ap
84	21	100.0	337	1	USG_ECOLI	P08390	escherichia
85	21	100.0	340	1	VIRS_MYCTU	Q06861	mycobacteri
86	21	100.0	341	1	YF43_MYCTU	Q10783	mycobacteri
87	21	100.0	342	1	RFC3_SCHPO	O14003	schizosacch
88	21	100.0	343	1	ARGC_METKA	Q8twf8	methanopyru
89	21	100.0	347	1	E13A_SOYBN	Q03773	glycine max
90	21	100.0	350	1	DCUP_SYNY3	P54224	synechocyst
91	21	100.0	350	1	OPSL_CALJA	P34989	callithrix
92	21	100.0	350	1	PTER_DROME	Q9vhf2	drosophila
93	21	100.0	353	1	ADA_SALTY	P26189	salmonella
94	21	100.0	359	1	OPSG_MOUSE	O35599	mus musculu
95	21	100.0	359	1	OPSG_RAT	O35476	rattus norv
96	21	100.0	364	1	FL3H_VITVI	P41090	vitis vinif
97	21	100.0	364	1	OPSG_CAVPO	Q9r024	cavia porce
98	21	100.0	364	1	OPSG_HUMAN	P04001	homo sapien
99	21	100.0	364	1	OPSG_RABIT	O18910	oryctolagus
100	21	100.0	364	1	OPSR_CAPHI	Q95170	capra hircu
101	21	100.0	364	1	OPSR_FELCA	O18913	felis silve
102	21	100.0	364	1	OPSR_HUMAN	P04000	homo sapien
103	21	100.0	367	1	DUS1_HUMAN	P28562	homo sapien
104	21	100.0	367	1	DUS1_MOUSE	P28563	mus musculu
105	21	100.0	367	1	DUS1_RAT	Q64623	rattus norv
106	21	100.0	368	1	AROB_SYNEL	Q8dks3	synechococc
107	21	100.0	372	1	MRF1_SCHPO	Q10488	schizosacch
108	21	100.0	373	1	YE52_LISMO	P53434	listeria mo
109	21	100.0	373	1	YE89_LISIN	Q92bq8	listeria in
110	21	100.0	377	1	NTRB_BRASR	P10578	bradyrhizob
111	21	100.0	380	1	GB12_HUMAN	Q03113	homo sapien
112	21	100.0	382	1	ISC1_ARCFU	O30052	archaeoglob
113	21	100.0	382	1	ISC2_ARCFU	O29689	archaeoglob
114	21	100.0	382	1	METX_CAUCR	Q9aas1	caulobacter
115	21	100.0	386	1	PGK_ECO57	Q8xd03	escherichia
116	21	100.0	386	1	PGK_ECOLI	P11665	escherichia
117	21	100.0	386	1	PGK_PHOLU	Q8gf87	photorhabdu
118	21	100.0	386	1	PGK_SALTY	Q8xg18	salmonella
119	21	100.0	386	1	PGK_VIBPA	Q871l1	vibrio para
120	21	100.0	386	1	PGK_VIBVU	Q8dca0	vibrio vuln
121	21	100.0	387	1	PGK_PSEAE	Q9i5y4	pseudomonas
122	21	100.0	387	1	PGK_PSEPK	Q88d64	pseudomonas
123	21	100.0	387	1	PGK_PSESM	Q88ak3	pseudomonas
124	21	100.0	387	1	PGK_VIBCH	P96154	vibrio chol
125	21	100.0	387	1	PGK_YERPE	Q8zhh3	yersinia pe
126	21	100.0	387	1	YJK7_YEAST	P42947	saccharomyc
127	21	100.0	388	1	PGK_BUCAP	Q8k9b3	buchnera ap
128	21	100.0	388	1	RFC5_CAEEL	P34429	caenorhabdi
129	21	100.0	389	1	PGK_THETH	P09403	thermus the
130	21	100.0	389	1	Y466_TREPA	O83479	treponema p
131	21	100.0	390	1	CNX2_ARATH	Q39055	arabidopsis

132	21	100.0	391	1	PGK_SHEON	Q8eib1 shewanella
133	21	100.0	392	1	PURT_PASHA	P46927 p probable
134	21	100.0	392	1	RD22_ARATH	Q08298 arabidopsis
135	21	100.0	392	1	YE62_MYCLE	Q49682 mycobacteri
136	21	100.0	393	1	PGK_BORBU	Q59181 borrelia bu
137	21	100.0	393	1	PGK_BUCBP	P59461 buchnera ap
138	21	100.0	393	1	PGK_OCEIH	Q8enp3 oceanobacil
139	21	100.0	394	1	PGK_BACHD	Q9k714 bacillus ha
140	21	100.0	394	1	PGK_BACME	P24269 bacillus me
141	21	100.0	394	1	PGK_BACST	P18912 bacillus st
142	21	100.0	394	1	PGK_BACSU	P40924 bacillus su
143	21	100.0	394	1	PGK_THETN	Q8r965 thermoanaer
144	21	100.0	394	1	PGK_WIGBR	Q8d2p9 wiggleswort
145	21	100.0	395	1	GMDS_DROME	Q9vmw9 drosophila
146	21	100.0	396	1	PGK_CAUCR	Q9a3f5 caulobacter
147	21	100.0	396	1	PGK_LEPIN	Q8f5h8 leptospira
148	21	100.0	397	1	THIM_HUMAN	P42765 homo sapien
149	21	100.0	397	1	THIM_RAT	P13437 rattus norv
150	21	100.0	400	1	CPXP_RHISN	P55544 rhizobium s

#### ALIGNMENTS

#### RESULT 1

#### RIPX\_CUCPE

ID RIPX\_CUCPE STANDARD; PRT; 20 AA.  
AC P80750;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Ribosome-inactivating protein (EC 3.2.2.22) (rRNA N-glycosidase)  
DE (Fragment).  
OS Cucurbita pepo (Vegetable marrow) (Summer squash).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.  
OX NCBI\_TaxID=3663;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Mesocarp;  
RX MEDLINE=97175025; PubMed=9022685;  
RA Yoshinari S., Yokota S., Sawmoto H., Koresawa S., Tamura M.,  
RA Endo Y.;  
RT "Purification, characterization and subcellular localization of a  
RT type-1 ribosome-inactivating protein from the sarcocarp of Cucurbita  
RT pepo.";  
RL Eur. J. Biochem. 242:585-591(1996).  
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of the N-glycosidic bond at one  
CC specific adenosine on the 28S rRNA.  
CC -!- SIMILARITY: Belongs to the ribosome-inactivating protein family.  
CC Type 1 RIP subfamily.  
DR InterPro; IPR001574; RIP.  
DR Pfam; PF00161; RIP; 1.  
DR PROSITE; PS00275; SHIGA\_RICIN; PARTIAL.  
KW Plant defense; Protein synthesis inhibitor; Hydrolase; Toxin.  
FT NON\_TER 20 20

SQ SEQUENCE 20 AA; 2235 MW; 8E753D217FDEAA6D CRC64;

Query Match 100.0%; Score 21; DB 1; Length 20;  
Best Local Similarity 100.0%; Pred. No. 7.4;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NVRF 4  
|||  
Db 1 NVRF 4

## RESULT 2

Y414\_HAEIN

ID Y414\_HAEIN STANDARD; PRT; 70 AA.

AC Q57392; O05021;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Hypothetical protein HI0414.

GN HI0414.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Haemophilus.

OX NCBI\_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

RA Venter J.C.;

RT "Whole-genome random sequencing and assembly of Haemophilus influenzae

RT Rd.";

RL Science 269:496-512(1995).

CC -!- SIMILARITY: STRONG, TO THE C-TERMINAL OF N.GONORRHOEAE OPACITY  
CC PROTEINS.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; U32725; AAC22081.1; -.

DR PIR; F64066; F64066.

DR TIGR; HI0414; -.

DR InterPro; IPR003394; Porin\_opacity.

DR Pfam; PF02462; Opacity; 1.

KW Hypothetical protein; Complete proteome.

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:09:21 ; Search time 16.1345 Seconds  
 (without alignments)  
 78.222 Million cell updates/sec

Title: US-10-030-735-51  
 Perfect score: 21  
 Sequence: 1 NVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertibrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%				
No.	Score	Match	Length	DB	ID	Description
-----						

1	21	100.0	24	4	Q8WYB9	Q8wyb9 homo sapien
2	21	100.0	28	16	Q8CLI3	Q8cli3 yersinia pe
3	21	100.0	39	16	Q8U557	Q8u557 agrobacteri
4	21	100.0	42	16	Q8KAE6	Q8kae6 chlorobium
5	21	100.0	44	6	Q9MYL7	Q9myl7 callithrix
6	21	100.0	44	6	Q9MZH6	Q9mzh6 callithrix
7	21	100.0	45	16	Q8EZ53	Q8ez53 leptospira
8	21	100.0	47	7	O98166	O98166 arctocephal
9	21	100.0	48	16	Q8EE67	Q8ee67 shewanella
10	21	100.0	55	6	Q8HZ39	Q8hz39 callimico g
11	21	100.0	55	6	Q8HZ37	Q8hz37 callimico g
12	21	100.0	55	6	Q8HZ35	Q8hz35 callithrix
13	21	100.0	55	6	Q8HZ24	Q8hz24 leontopithe
14	21	100.0	56	7	Q99938	Q99938 homo sapien
15	21	100.0	57	7	Q860Q3	Q860q3 microcebus
16	21	100.0	61	12	Q91F79	Q91f79 chilo iride
17	21	100.0	62	16	Q8A9N2	Q8a9n2 bacteroides
18	21	100.0	68	12	Q91F78	Q91f78 chilo iride
19	21	100.0	68	12	Q91F80	Q91f80 chilo iride
20	21	100.0	68	12	Q91F77	Q91f77 chilo iride
21	21	100.0	68	12	Q91F81	Q91f81 chilo iride
22	21	100.0	69	7	Q30127	Q30127 homo sapien
23	21	100.0	69	7	Q30409	Q30409 cervus elap
24	21	100.0	69	7	Q95HI5	Q95hi5 canis lupus
25	21	100.0	69	7	Q95HH8	Q95hh8 canis latra
26	21	100.0	69	7	Q95HG8	Q95hg8 canis latra
27	21	100.0	69	7	Q95HG9	Q95hg9 canis latra
28	21	100.0	69	7	Q95HH3	Q95hh3 canis latra
29	21	100.0	69	7	Q95HI1	Q95hi1 canis lupus
30	21	100.0	72	2	Q8KMZ4	Q8kmz4 escherichia
31	21	100.0	72	7	Q29648	Q29648 homo sapien
32	21	100.0	73	7	Q31644	Q31644 homo sapien
33	21	100.0	73	7	Q29649	Q29649 homo sapien
34	21	100.0	73	7	Q95498	Q95498 pan paniscu
35	21	100.0	73	7	Q95499	Q95499 pan paniscu
36	21	100.0	73	7	O19723	O19723 homo sapien
37	21	100.0	73	7	Q31639	Q31639 homo sapien
38	21	100.0	73	7	Q29647	Q29647 homo sapien
39	21	100.0	74	6	Q9TTM1	Q9ttm1 bos taurus
40	21	100.0	74	6	Q9TTM0	Q9ttm0 bos taurus
41	21	100.0	74	7	Q9MY00	Q9my00 homo sapien
42	21	100.0	74	7	O19696	O19696 homo sapien
43	21	100.0	74	7	Q9MY16	Q9my16 homo sapien
44	21	100.0	74	7	Q29903	Q29903 homo sapien
45	21	100.0	74	7	Q9MY17	Q9my17 homo sapien
46	21	100.0	74	7	Q29804	Q29804 homo sapien
47	21	100.0	74	7	Q9MXZ7	Q9mxz7 homo sapien
48	21	100.0	74	7	Q9MXZ8	Q9mxz8 homo sapien
49	21	100.0	75	2	Q93EZ3	Q93ez3 shigella fl
50	21	100.0	75	6	Q9TTL8	Q9ttl8 bos taurus
51	21	100.0	75	7	Q29687	Q29687 homo sapien
52	21	100.0	75	7	Q8MGY2	Q8mgy2 homo sapien
53	21	100.0	75	7	Q29686	Q29686 homo sapien
54	21	100.0	75	7	Q8MGY3	Q8mgy3 homo sapien
55	21	100.0	75	7	Q7YPW3	Q7ypw3 homo sapien
56	21	100.0	75	16	Q8YSR5	Q8ysr5 anabaena sp
57	21	100.0	75	17	Q97C24	Q97c24 thermoplasm

58	21	100.0	76	7	O46816	O46816 equus asinu
59	21	100.0	76	7	Q9TPC2	Q9tpc2 macaca mula
60	21	100.0	77	7	Q9GJ82	Q9gj82 ovis aries
61	21	100.0	77	7	Q9GJ91	Q9gj91 ovis aries
62	21	100.0	77	11	O89067	O89067 mus musculu
63	21	100.0	78	6	O97839	O97839 ovis aries
64	21	100.0	78	6	O97838	O97838 ovis aries
65	21	100.0	78	7	O19592	O19592 homo sapien
66	21	100.0	78	7	Q9TP03	Q9tp03 homo sapien
67	21	100.0	78	7	Q8MHI7	Q8mhi7 bos indicus
68	21	100.0	78	7	Q8MHI0	Q8mhi0 bos indicus
69	21	100.0	78	7	Q30293	Q30293 bos taurus
70	21	100.0	78	7	Q29906	Q29906 homo sapien
71	21	100.0	78	7	Q07136	Q07136 bos taurus
72	21	100.0	78	7	Q8MHJ9	Q8mhj9 bubalus bub
73	21	100.0	78	7	O19594	O19594 homo sapien
74	21	100.0	78	7	Q9TQ23	Q9tq23 homo sapien
75	21	100.0	78	7	Q861B8	Q861b8 homo sapien
76	21	100.0	78	7	Q860S1	Q860s1 homo sapien
77	21	100.0	78	7	Q860B8	Q860b8 homo sapien
78	21	100.0	78	7	Q7YP63	Q7yp63 homo sapien
79	21	100.0	78	16	Q7U081	Q7u081 mycobacteri
80	21	100.0	79	7	Q9TPW7	Q9tpw7 equus cabal
81	21	100.0	79	7	O46698	O46698 homo sapien
82	21	100.0	80	2	Q8KJW5	Q8kjlw5 proteus vul
83	21	100.0	80	7	O02969	O02969 homo sapien
84	21	100.0	80	7	Q29806	Q29806 homo sapien
85	21	100.0	80	7	O77970	O77970 homo sapien
86	21	100.0	80	7	O02951	O02951 homo sapien
87	21	100.0	81	6	O77825	O77825 bos taurus
88	21	100.0	81	7	Q29887	Q29887 homo sapien
89	21	100.0	81	7	O19791	O19791 homo sapien
90	21	100.0	81	7	Q9GIL4	Q9gil4 homo sapien
91	21	100.0	81	7	P79446	P79446 bos taurus
92	21	100.0	81	7	Q30740	Q30740 macaca neme
93	21	100.0	81	7	Q9TPW8	Q9tpw8 homo sapien
94	21	100.0	81	7	P79445	P79445 bos taurus
95	21	100.0	81	7	P79448	P79448 bos taurus
96	21	100.0	81	7	Q8WLV6	Q8wlv6 bos taurus
97	21	100.0	81	7	Q95382	Q95382 homo sapien
98	21	100.0	81	7	Q30739	Q30739 macaca neme
99	21	100.0	81	7	Q30563	Q30563 macaca fasc
100	21	100.0	81	7	Q8HWK4	Q8hkw4 bubalus bub
101	21	100.0	81	7	Q95541	Q95541 bos taurus
102	21	100.0	82	5	Q19078	Q19078 caenorhabdi
103	21	100.0	82	7	O77965	O77965 homo sapien
104	21	100.0	82	7	Q30639	Q30639 macaca mula
105	21	100.0	82	7	Q9MXN2	Q9mxn2 macaca mula
106	21	100.0	82	7	Q30637	Q30637 macaca mula
107	21	100.0	82	7	Q30615	Q30615 macaca mula
108	21	100.0	82	7	Q30602	Q30602 macaca mula
109	21	100.0	82	7	Q30638	Q30638 macaca mula
110	21	100.0	82	7	Q30636	Q30636 macaca mula
111	21	100.0	82	7	Q30965	Q30965 pan troglod
112	21	100.0	82	7	Q29818	Q29818 homo sapien
113	21	100.0	82	7	Q30966	Q30966 pan troglod
114	21	100.0	82	7	Q7YQ16	Q7yq16 macaca mula

115	21	100.0	83	4	Q15945	Q15945 homo sapien
116	21	100.0	83	4	Q14280	Q14280 homo sapien
117	21	100.0	83	6	Q9TSS5	Q9tss5 bos indicus
118	21	100.0	83	7	Q9BDA2	Q9bda2 ovis canade
119	21	100.0	83	7	077978	077978 homo sapien
120	21	100.0	83	7	062890	062890 homo sapien
121	21	100.0	83	7	Q9GIU0	Q9giu0 damaliscus
122	21	100.0	83	7	Q29955	Q29955 homo sapien
123	21	100.0	83	7	Q29725	Q29725 homo sapien
124	21	100.0	83	7	Q9GIU9	Q9giu9 damaliscus
125	21	100.0	83	7	Q9BD95	Q9bd95 ovis canade
126	21	100.0	83	7	Q9BDA5	Q9bda5 ovis canade
127	21	100.0	83	7	Q9BD93	Q9bd93 ovis canade
128	21	100.0	83	7	Q9GIU1	Q9giu1 damaliscus
129	21	100.0	83	7	Q9BD90	Q9bd90 ovis canade
130	21	100.0	83	7	Q30276	Q30276 bos taurus
131	21	100.0	83	7	Q9BD91	Q9bd91 ovis canade
132	21	100.0	83	7	Q9GIU8	Q9giu8 damaliscus
133	21	100.0	83	7	Q9BD89	Q9bd89 ovis canade
134	21	100.0	83	7	P79550	P79550 homo sapien
135	21	100.0	83	7	Q30275	Q30275 bos taurus
136	21	100.0	83	7	Q9GIU3	Q9giu3 damaliscus
137	21	100.0	83	7	Q9GIJ8	Q9gij8 damaliscus
138	21	100.0	83	7	Q9GIK0	Q9gik0 damaliscus
139	21	100.0	84	7	Q29706	Q29706 homo sapien
140	21	100.0	84	7	O02876	O02876 homo sapien
141	21	100.0	84	7	Q8HWN4	Q8hwn4 homo sapien
142	21	100.0	85	7	Q30199	Q30199 homo sapien
143	21	100.0	85	7	Q30111	Q30111 homo sapien
144	21	100.0	85	7	078098	078098 bos taurus
145	21	100.0	85	7	Q9Y4H8	Q9y4h8 homo sapien
146	21	100.0	85	7	Q30811	Q30811 ovis aries
147	21	100.0	85	7	Q29711	Q29711 homo sapien
148	21	100.0	85	7	Q29833	Q29833 homo sapien
149	21	100.0	85	7	Q8MHD8	Q8mhd8 peromyscus
150	21	100.0	85	7	Q8HWN2	Q8hwn2 homo sapien

# ALIGNMENTS

## RESULT 1

Q8WYB9

ID Q8WYB9 PRELIMINARY; PRT; 24 AA.

AC Q8WYB9;

DT 01-MAR-2002 (TrEMBLrel. 20, Created)

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)

DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)

DE Dystrophin (Fragment).

GN DMD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE OF 7-24 FROM N.A.

RX MEDLINE=97107974; PubMed=8950674;



RA Roest P.A., Bout M., van der Tuijn A.C., Ginjaar I.B., Bakker E.,  
 RA Hogervorst F.B., van Ommen G.J., den Dunnen J.T.;  
 RT "Splicing mutations in DMD/BMD detected by RT-PCR/PTT: detection of a  
 RT 19AA insertion in the cysteine rich domain of dystrophin compatible  
 RT with BMD.";  
 RL J. Med. Genet. 33:935-939(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Roberts R.G.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA den Dunnen J.T.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF213441; AAL61585.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2868 MW; 9C932A284F14E382 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 96;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NVRF 4  
 ||||  
 Db 4 NVRF 7

# RESULT 2

## Q8CLI3

ID Q8CLI3 PRELIMINARY; PRT; 28 AA.  
 AC Q8CLI3;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
 DE Hypothetical.  
 GN Y1274.  
 OS Yersinia pestis.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Yersinia.  
 OX NCBI\_TaxID=632;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=KIM5 / Biovar Mediaevalis;  
 RX MEDLINE=22137863; PubMed=12142430;  
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,  
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,  
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,  
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,  
 RA Perry R.D.;  
 RT "Genome sequence of Yersinia pestis KIM.";  
 RL J. Bacteriol. 184:4601-4611(2002).  
 DR EMBL; AE013730; AAM84848.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 28 AA; 3113 MW; 07B7A00783BC2F68 CRC64;

Query Match 100.0%; Score 21; DB 16; Length 28;

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:16:36 ; Search time 6.28571 Seconds  
 (without alignments)  
 32.853 Million cell updates/sec

Title: US-10-030-735-52  
 Perfect score: 19  
 Sequence: 1 SVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	19	100.0	15	4	US-09-009-953-143	Sequence 143, App
2	19	100.0	20	4	US-09-556-877-154	Sequence 154, App
3	19	100.0	20	4	US-09-556-877-155	Sequence 155, App
4	19	100.0	20	4	US-09-620-412C-154	Sequence 154, App
5	19	100.0	20	4	US-09-620-412C-155	Sequence 155, App
6	19	100.0	20	4	US-09-598-419-154	Sequence 154, App
7	19	100.0	20	4	US-09-598-419-155	Sequence 155, App
8	19	100.0	22	2	US-08-480-190-43	Sequence 43, Appl
9	19	100.0	22	2	US-08-488-379-43	Sequence 43, Appl
10	19	100.0	22	4	US-08-475-399A-43	Sequence 43, Appl
11	19	100.0	22	5	PCT-US93-07545-43	Sequence 43, Appl

12	19	100.0	31	4	US-08-525-539A-13	Sequence 13, Appl
13	19	100.0	32	4	US-08-525-539A-10	Sequence 10, Appl
14	19	100.0	32	4	US-08-525-539A-14	Sequence 14, Appl
15	19	100.0	34	4	US-08-525-539A-16	Sequence 16, Appl
16	19	100.0	36	1	US-08-039-137-38	Sequence 38, Appl
17	19	100.0	36	1	US-08-039-137-39	Sequence 39, Appl
18	19	100.0	40	6	5196194-9	Patent No. 5196194
19	19	100.0	52	1	US-08-159-340A-7	Sequence 7, Appli
20	19	100.0	53	4	US-09-556-877-157	Sequence 157, App
21	19	100.0	53	4	US-09-620-412C-157	Sequence 157, App
22	19	100.0	53	4	US-09-598-419-157	Sequence 157, App
23	19	100.0	62	4	US-08-311-731A-328	Sequence 328, App
24	19	100.0	68	3	US-09-120-887-2	Sequence 2, Appli
25	19	100.0	73	4	US-09-333-809-10	Sequence 10, Appl
26	19	100.0	73	4	US-09-333-809-59	Sequence 59, Appl
27	19	100.0	78	4	US-09-333-809-21	Sequence 21, Appl
28	19	100.0	79	4	US-09-333-809-15	Sequence 15, Appl
29	19	100.0	79	4	US-09-333-809-58	Sequence 58, Appl
30	19	100.0	80	4	US-09-333-809-19	Sequence 19, Appl
31	19	100.0	80	4	US-09-333-809-22	Sequence 22, Appl
32	19	100.0	80	4	US-09-333-809-30	Sequence 30, Appl
33	19	100.0	80	4	US-09-333-809-91	Sequence 91, Appl
34	19	100.0	81	4	US-09-333-809-4	Sequence 4, Appli
35	19	100.0	81	4	US-09-333-809-11	Sequence 11, Appl
36	19	100.0	81	4	US-09-333-809-99	Sequence 99, Appl
37	19	100.0	82	2	US-08-248-839C-122	Sequence 122, App
38	19	100.0	83	4	US-09-333-809-60	Sequence 60, Appl
39	19	100.0	85	4	US-09-333-809-13	Sequence 13, Appl
40	19	100.0	85	4	US-09-333-809-17	Sequence 17, Appl
41	19	100.0	86	4	US-09-333-809-23	Sequence 23, Appl
42	19	100.0	87	3	US-08-905-223-334	Sequence 334, App
43	19	100.0	87	4	US-09-333-809-14	Sequence 14, Appl
44	19	100.0	88	4	US-09-333-809-8	Sequence 8, Appli
45	19	100.0	89	1	US-08-039-137-17	Sequence 17, Appl
46	19	100.0	90	4	US-09-489-039A-9352	Sequence 9352, Ap
47	19	100.0	94	3	US-08-960-190A-37	Sequence 37, Appl
48	19	100.0	97	4	US-09-252-991A-25508	Sequence 25508, A
49	19	100.0	98	4	US-09-333-809-12	Sequence 12, Appl
50	19	100.0	98	4	US-09-540-236-2245	Sequence 2245, Ap
51	19	100.0	102	4	US-09-333-809-2	Sequence 2, Appli
52	19	100.0	102	4	US-09-333-809-3	Sequence 3, Appli
53	19	100.0	102	4	US-09-333-809-5	Sequence 5, Appli
54	19	100.0	102	4	US-09-333-809-6	Sequence 6, Appli
55	19	100.0	102	4	US-09-333-809-7	Sequence 7, Appli
56	19	100.0	102	4	US-09-333-809-9	Sequence 9, Appli
57	19	100.0	102	4	US-09-333-809-16	Sequence 16, Appl
58	19	100.0	102	4	US-09-333-809-18	Sequence 18, Appl
59	19	100.0	102	4	US-09-333-809-20	Sequence 20, Appl
60	19	100.0	102	4	US-09-333-809-45	Sequence 45, Appl
61	19	100.0	106	4	US-09-107-532A-4765	Sequence 4765, Ap
62	19	100.0	108	4	US-09-199-637A-227	Sequence 227, App
63	19	100.0	113	4	US-09-107-532A-4226	Sequence 4226, Ap
64	19	100.0	115	1	US-08-052-681-3	Sequence 3, Appli
65	19	100.0	116	1	US-08-052-681-4	Sequence 4, Appli
66	19	100.0	118	2	US-08-846-762-10	Sequence 10, Appl
67	19	100.0	124	4	US-09-700-847-6	Sequence 6, Appli
68	19	100.0	127	4	US-09-252-991A-25275	Sequence 25275, A

69	19	100.0	130	4	US-09-107-532A-6103	Sequence 6103, Ap
70	19	100.0	136	4	US-09-107-532A-6354	Sequence 6354, Ap
71	19	100.0	138	4	US-09-198-452A-471	Sequence 471, App
72	19	100.0	140	3	US-09-179-558-60	Sequence 60, Appl
73	19	100.0	140	4	US-09-722-825-60	Sequence 60, Appl
74	19	100.0	140	4	US-09-722-487-60	Sequence 60, Appl
75	19	100.0	140	4	US-09-722-708-60	Sequence 60, Appl
76	19	100.0	150	4	US-09-252-991A-28955	Sequence 28955, A
77	19	100.0	152	4	US-09-252-991A-18069	Sequence 18069, A
78	19	100.0	154	3	US-09-193-104-7	Sequence 7, Appli
79	19	100.0	154	3	US-09-193-104-13	Sequence 13, Appl
80	19	100.0	154	3	US-09-193-104-14	Sequence 14, Appl
81	19	100.0	154	3	US-09-193-104-15	Sequence 15, Appl
82	19	100.0	154	3	US-09-193-104-16	Sequence 16, Appl
83	19	100.0	154	3	US-09-193-104-17	Sequence 17, Appl
84	19	100.0	154	3	US-09-193-104-18	Sequence 18, Appl
85	19	100.0	154	3	US-09-193-104-19	Sequence 19, Appl
86	19	100.0	154	3	US-09-193-104-20	Sequence 20, Appl
87	19	100.0	154	3	US-09-193-104-21	Sequence 21, Appl
88	19	100.0	154	3	US-09-193-104-22	Sequence 22, Appl
89	19	100.0	154	3	US-09-193-104-23	Sequence 23, Appl
90	19	100.0	154	3	US-09-193-104-24	Sequence 24, Appl
91	19	100.0	154	3	US-09-193-104-25	Sequence 25, Appl
92	19	100.0	157	4	US-09-306-420C-20	Sequence 20, Appl
93	19	100.0	158	4	US-09-489-039A-8612	Sequence 8612, Ap
94	19	100.0	165	4	US-09-252-991A-24560	Sequence 24560, A
95	19	100.0	168	4	US-09-148-545-184	Sequence 184, App
96	19	100.0	169	4	US-09-540-236-2163	Sequence 2163, Ap
97	19	100.0	174	3	US-09-153-586-22	Sequence 22, Appl
98	19	100.0	178	4	US-09-107-532A-3762	Sequence 3762, Ap
99	19	100.0	179	2	US-08-993-228-33	Sequence 33, Appl
100	19	100.0	179	4	US-09-389-681-179	Sequence 179, App
101	19	100.0	179	4	US-09-620-405B-179	Sequence 179, App
102	19	100.0	179	4	US-09-339-338-179	Sequence 179, App
103	19	100.0	179	4	US-09-433-826B-179	Sequence 179, App
104	19	100.0	179	4	US-09-604-287A-179	Sequence 179, App
105	19	100.0	179	4	US-09-285-480-179	Sequence 179, App
106	19	100.0	179	4	US-09-834-759-179	Sequence 179, App
107	19	100.0	184	4	US-09-543-681A-8152	Sequence 8152, Ap
108	19	100.0	191	4	US-09-134-001C-4710	Sequence 4710, Ap
109	19	100.0	191	4	US-09-252-991A-18626	Sequence 18626, A
110	19	100.0	194	4	US-09-252-991A-21846	Sequence 21846, A
111	19	100.0	216	4	US-09-556-877-20	Sequence 20, Appl
112	19	100.0	216	4	US-09-288-594A-20	Sequence 20, Appl
113	19	100.0	216	4	US-09-620-412C-20	Sequence 20, Appl
114	19	100.0	216	4	US-09-410-568-20	Sequence 20, Appl
115	19	100.0	216	4	US-09-598-419-20	Sequence 20, Appl
116	19	100.0	225	6	5436139-4	Patent No. 5436139
117	19	100.0	226	1	US-08-378-011A-3	Sequence 3, Appli
118	19	100.0	226	4	US-09-471-573A-2	Sequence 2, Appli
119	19	100.0	226	4	US-09-471-573A-40	Sequence 40, Appl
120	19	100.0	226	6	5196194-21	Patent No. 5196194
121	19	100.0	226	6	5198348-1	Patent No. 5198348
122	19	100.0	226	6	5436139-5	Patent No. 5436139
123	19	100.0	228	1	US-08-447-591-2	Sequence 2, Appli
124	19	100.0	228	1	US-08-447-591-3	Sequence 3, Appli
125	19	100.0	228	1	US-08-450-943-2	Sequence 2, Appli

126	19	100.0	228	1	US-08-450-943-3	Sequence 3, Appli
127	19	100.0	228	1	US-08-059-031-2	Sequence 2, Appli
128	19	100.0	228	1	US-08-059-031-3	Sequence 3, Appli
129	19	100.0	228	1	US-08-167-035-47	Sequence 47, Appl
130	19	100.0	228	1	US-08-208-887A-47	Sequence 47, Appl
131	19	100.0	228	2	US-08-539-005-47	Sequence 47, Appl
132	19	100.0	228	2	US-08-815-176-5	Sequence 5, Appli
133	19	100.0	228	2	US-08-450-942-2	Sequence 2, Appli
134	19	100.0	228	2	US-08-450-942-3	Sequence 3, Appli
135	19	100.0	228	4	US-09-280-598-44	Sequence 44, Appl
136	19	100.0	228	4	US-09-197-344-5	Sequence 5, Appli
137	19	100.0	228	5	PCT-US94-05090-2	Sequence 2, Appli
138	19	100.0	228	5	PCT-US94-05090-3	Sequence 3, Appli
139	19	100.0	230	4	US-09-489-039A-13797	Sequence 13797, A
140	19	100.0	235	4	US-09-252-991A-28794	Sequence 28794, A
141	19	100.0	236	1	US-08-378-011A-1	Sequence 1, Appli
142	19	100.0	240	4	US-09-252-991A-28485	Sequence 28485, A
143	19	100.0	242	4	US-09-107-532A-6951	Sequence 6951, Ap
144	19	100.0	244	3	US-08-463-903-24	Sequence 24, Appl
145	19	100.0	244	4	US-07-935-695-24	Sequence 24, Appl
146	19	100.0	245	4	US-09-252-991A-32951	Sequence 32951, A
147	19	100.0	245	4	US-09-198-452A-1091	Sequence 1091, Ap
148	19	100.0	248	4	US-09-134-001C-5271	Sequence 5271, Ap
149	19	100.0	253	4	US-09-540-236-2592	Sequence 2592, Ap
150	19	100.0	254	4	US-09-252-991A-30634	Sequence 30634, A

#### ALIGNMENTS

#### RESULT 1

US-09-009-953-143

; Sequence 143, Application US/09009953

; Patent No. 6413517

; GENERAL INFORMATION:

; APPLICANT: Sette, Alessandro

; TITLE OF INVENTION: Identification of Broadly

; Reactive DR Restricted Epitopes

; NUMBER OF SEQUENCES: 274

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend and Crew LLP

; STREET: Two Embarcadero Center, Eighth Floor

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/009,953

; FILING DATE: 21-Jan-1998

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/036,713

```

; FILING DATE: 23-JAN-1997
; APPLICATION NUMBER: US 60/037,432
; FILING DATE: 07-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Weber, Ellen Lauver
; REGISTRATION NUMBER: 32,762
; REFERENCE/DOCKET NUMBER: 018623-011520US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-576-0200
; TELEFAX: 415-576-0300
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 143:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 143:
US-09-009-953-143

```

```

Query Match          100.0%;  Score 19;  DB 4;  Length 15;
Best Local Similarity 100.0%;  Pred. No. 51;
Matches      4;  Conservative    0;  Mismatches    0;  Indels      0;  Gaps      0;

```

```

Qy      1 SVRF 4
        ||||
Db      1 SVRF 4

```

# RESULT 2

```

US-09-556-877-154
; Sequence 154, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF CHLAMYDIAL INFECTION
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 154
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
US-09-556-877-154

```

```

Query Match          100.0%;  Score 19;  DB 4;  Length 20;
Best Local Similarity 100.0%;  Pred. No. 67;

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 22:05:26 ; Search time 16.8403 Seconds  
 (without alignments)  
 62.382 Million cell updates/sec

Title: US-10-030-735-52  
 Perfect score: 19  
 Sequence: 1 SVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Query	%				
No.	Score	Match	Length	DB	ID	Description

1	19	100.0	8	9	US-09-169-048-4	Sequence 4, Appli
2	19	100.0	8	10	US-09-997-209-4	Sequence 4, Appli
3	19	100.0	8	10	US-09-839-469-4	Sequence 4, Appli
4	19	100.0	9	14	US-10-283-423-115	Sequence 115, App
5	19	100.0	9	14	US-10-213-821-115	Sequence 115, App
6	19	100.0	9	16	US-10-365-761B-43	Sequence 43, Appl
7	19	100.0	16	14	US-10-225-567A-1872	Sequence 1872, Ap
8	19	100.0	20	8	US-08-677-599B-5	Sequence 5, Appli
9	19	100.0	20	9	US-09-864-761-38469	Sequence 38469, A
10	19	100.0	20	9	US-09-841-132-154	Sequence 154, App
11	19	100.0	20	9	US-09-841-132-155	Sequence 155, App
12	19	100.0	22	8	US-08-677-599B-4	Sequence 4, Appli
13	19	100.0	25	12	US-10-424-599-208573	Sequence 208573,
14	19	100.0	31	9	US-09-956-206A-13	Sequence 13, Appl
15	19	100.0	32	9	US-09-956-206A-10	Sequence 10, Appl
16	19	100.0	32	9	US-09-956-206A-14	Sequence 14, Appl
17	19	100.0	34	9	US-09-956-206A-16	Sequence 16, Appl
18	19	100.0	37	12	US-10-424-599-150570	Sequence 150570,
19	19	100.0	38	12	US-10-424-599-241220	Sequence 241220,
20	19	100.0	39	12	US-10-424-599-276455	Sequence 276455,
21	19	100.0	41	12	US-10-424-599-166083	Sequence 166083,
22	19	100.0	41	12	US-10-424-599-214476	Sequence 214476,
23	19	100.0	42	9	US-09-975-143-10	Sequence 10, Appl
24	19	100.0	42	12	US-10-424-599-166277	Sequence 166277,
25	19	100.0	44	12	US-10-424-599-213832	Sequence 213832,
26	19	100.0	44	12	US-10-424-599-235615	Sequence 235615,
27	19	100.0	45	9	US-09-888-615-116	Sequence 116, App
28	19	100.0	45	12	US-10-424-599-253267	Sequence 253267,
29	19	100.0	47	10	US-09-948-820-81	Sequence 81, Appl
30	19	100.0	47	12	US-10-424-599-256103	Sequence 256103,
31	19	100.0	49	12	US-10-424-599-262031	Sequence 262031,
32	19	100.0	51	12	US-10-424-599-161526	Sequence 161526,
33	19	100.0	51	12	US-10-424-599-180514	Sequence 180514,
34	19	100.0	52	12	US-10-424-599-212054	Sequence 212054,
35	19	100.0	52	12	US-10-424-599-256950	Sequence 256950,
36	19	100.0	52	14	US-10-106-698-8291	Sequence 8291, Ap
37	19	100.0	53	9	US-09-841-132-157	Sequence 157, App
38	19	100.0	54	12	US-10-425-114-42841	Sequence 42841, A
39	19	100.0	55	12	US-10-424-599-250575	Sequence 250575,
40	19	100.0	55	12	US-10-424-599-269119	Sequence 269119,
41	19	100.0	55	14	US-10-029-386-28389	Sequence 28389, A
42	19	100.0	56	12	US-10-424-599-163386	Sequence 163386,
43	19	100.0	57	12	US-10-424-599-267624	Sequence 267624,
44	19	100.0	58	12	US-10-424-599-179106	Sequence 179106,
45	19	100.0	58	12	US-10-424-599-254574	Sequence 254574,
46	19	100.0	60	9	US-09-925-299-890	Sequence 890, App
47	19	100.0	60	10	US-09-925-299-890	Sequence 890, App
48	19	100.0	60	12	US-10-424-599-220145	Sequence 220145,
49	19	100.0	60	14	US-10-106-698-4558	Sequence 4558, Ap
50	19	100.0	61	14	US-10-106-698-6110	Sequence 6110, Ap
51	19	100.0	64	12	US-10-424-599-213618	Sequence 213618,
52	19	100.0	64	12	US-10-424-599-253993	Sequence 253993,
53	19	100.0	64	12	US-10-424-599-280575	Sequence 280575,
54	19	100.0	65	12	US-10-424-599-249685	Sequence 249685,
55	19	100.0	66	12	US-10-424-599-153262	Sequence 153262,
56	19	100.0	66	12	US-10-424-599-219749	Sequence 219749,



57	19	100.0	66	12	US-10-424-599-224463	Sequence 224463,
58	19	100.0	67	12	US-10-424-599-282734	Sequence 282734,
59	19	100.0	69	9	US-09-764-860-438	Sequence 438, App
60	19	100.0	69	11	US-09-864-408A-2498	Sequence 2498, Ap
61	19	100.0	69	14	US-10-074-095-438	Sequence 438, App
62	19	100.0	69	15	US-10-212-872-438	Sequence 438, App
63	19	100.0	70	11	US-09-864-408A-4498	Sequence 4498, Ap
64	19	100.0	71	12	US-10-424-599-215425	Sequence 215425,
65	19	100.0	72	12	US-10-424-599-222551	Sequence 222551,
66	19	100.0	72	12	US-10-424-599-227469	Sequence 227469,
67	19	100.0	72	12	US-10-425-114-53592	Sequence 53592, A
68	19	100.0	73	9	US-09-867-550-1932	Sequence 1932, Ap
69	19	100.0	73	11	US-09-833-245-1320	Sequence 1320, Ap
70	19	100.0	73	14	US-10-156-761-7993	Sequence 7993, Ap
71	19	100.0	74	12	US-10-424-599-154888	Sequence 154888,
72	19	100.0	74	12	US-10-424-599-230124	Sequence 230124,
73	19	100.0	75	12	US-10-424-599-162950	Sequence 162950,
74	19	100.0	75	12	US-10-424-599-247799	Sequence 247799,
75	19	100.0	75	12	US-10-424-599-257099	Sequence 257099,
76	19	100.0	76	12	US-10-424-599-224517	Sequence 224517,
77	19	100.0	78	11	US-09-864-408A-3962	Sequence 3962, Ap
78	19	100.0	78	12	US-10-424-599-238324	Sequence 238324,
79	19	100.0	78	14	US-10-080-170-221	Sequence 221, App
80	19	100.0	78	14	US-10-080-170-355	Sequence 355, App
81	19	100.0	79	12	US-10-424-599-262182	Sequence 262182,
82	19	100.0	79	14	US-10-157-031-329	Sequence 329, App
83	19	100.0	80	12	US-10-424-599-172045	Sequence 172045,
84	19	100.0	80	12	US-10-424-599-173985	Sequence 173985,
85	19	100.0	80	12	US-10-424-599-265643	Sequence 265643,
86	19	100.0	81	12	US-10-424-599-237620	Sequence 237620,
87	19	100.0	82	12	US-10-424-599-188967	Sequence 188967,
88	19	100.0	82	12	US-10-424-599-259830	Sequence 259830,
89	19	100.0	82	12	US-10-424-599-265637	Sequence 265637,
90	19	100.0	83	12	US-10-424-599-199072	Sequence 199072,
91	19	100.0	85	12	US-10-424-599-215706	Sequence 215706,
92	19	100.0	85	12	US-10-424-599-269157	Sequence 269157,
93	19	100.0	85	14	US-10-106-698-6956	Sequence 6956, Ap
94	19	100.0	86	12	US-10-424-599-248096	Sequence 248096,
95	19	100.0	86	14	US-10-106-698-6502	Sequence 6502, Ap
96	19	100.0	88	12	US-10-424-599-150626	Sequence 150626,
97	19	100.0	89	12	US-10-424-599-215813	Sequence 215813,
98	19	100.0	89	12	US-10-424-599-247310	Sequence 247310,
99	19	100.0	89	12	US-10-424-599-248337	Sequence 248337,
100	19	100.0	90	12	US-10-424-599-165221	Sequence 165221,
101	19	100.0	90	12	US-10-424-599-284566	Sequence 284566,
102	19	100.0	91	12	US-10-424-599-248098	Sequence 248098,
103	19	100.0	93	9	US-09-764-877-1262	Sequence 1262, Ap
104	19	100.0	93	13	US-10-081-281-121	Sequence 121, App
105	19	100.0	93	15	US-10-242-515-1262	Sequence 1262, Ap
106	19	100.0	94	9	US-09-766-378A-37	Sequence 37, Appl
107	19	100.0	94	12	US-10-424-599-273018	Sequence 273018,
108	19	100.0	95	12	US-10-424-599-186606	Sequence 186606,
109	19	100.0	96	10	US-09-013-077A-15	Sequence 15, Appl
110	19	100.0	96	12	US-10-424-599-185817	Sequence 185817,
111	19	100.0	96	12	US-10-424-599-217459	Sequence 217459,
112	19	100.0	97	12	US-10-424-599-187160	Sequence 187160,
113	19	100.0	97	12	US-10-424-599-196064	Sequence 196064,

114	19	100.0	97	12	US-10-424-599-272721	Sequence 272721,
115	19	100.0	98	12	US-10-424-599-145426	Sequence 145426,
116	19	100.0	98	12	US-10-424-599-252401	Sequence 252401,
117	19	100.0	102	12	US-10-424-599-263081	Sequence 263081,
118	19	100.0	103	12	US-10-424-599-164093	Sequence 164093,
119	19	100.0	103	15	US-10-108-260A-3765	Sequence 3765, Ap
120	19	100.0	105	15	US-10-104-047-2457	Sequence 2457, Ap
121	19	100.0	105	15	US-10-108-260A-2680	Sequence 2680, Ap
122	19	100.0	106	12	US-10-424-599-175630	Sequence 175630,
123	19	100.0	106	12	US-10-424-599-231274	Sequence 231274,
124	19	100.0	107	12	US-10-424-599-214187	Sequence 214187,
125	19	100.0	108	10	US-09-975-719-227	Sequence 227, App
126	19	100.0	108	12	US-10-424-599-249291	Sequence 249291,
127	19	100.0	110	12	US-10-424-599-213313	Sequence 213313,
128	19	100.0	110	12	US-10-335-977-7703	Sequence 7703, Ap
129	19	100.0	110	14	US-10-083-357-920	Sequence 920, App
130	19	100.0	111	12	US-10-335-977-7702	Sequence 7702, Ap
131	19	100.0	112	12	US-10-424-599-147898	Sequence 147898,
132	19	100.0	113	12	US-10-424-599-251700	Sequence 251700,
133	19	100.0	114	12	US-10-424-599-194545	Sequence 194545,
134	19	100.0	114	15	US-10-108-260A-3998	Sequence 3998, Ap
135	19	100.0	116	12	US-10-424-599-202136	Sequence 202136,
136	19	100.0	116	12	US-10-424-599-235583	Sequence 235583,
137	19	100.0	117	12	US-10-424-599-204887	Sequence 204887,
138	19	100.0	117	12	US-10-424-599-258833	Sequence 258833,
139	19	100.0	118	12	US-10-425-114-41144	Sequence 41144, A
140	19	100.0	118	14	US-10-216-209-10	Sequence 10, Appl
141	19	100.0	120	12	US-10-425-114-41491	Sequence 41491, A
142	19	100.0	121	12	US-10-424-599-161821	Sequence 161821,
143	19	100.0	121	12	US-10-424-599-225585	Sequence 225585,
144	19	100.0	121	12	US-10-335-977-9643	Sequence 9643, Ap
145	19	100.0	122	12	US-10-424-599-190123	Sequence 190123,
146	19	100.0	122	12	US-10-424-599-263412	Sequence 263412,
147	19	100.0	124	12	US-10-424-599-263921	Sequence 263921,
148	19	100.0	124	15	US-10-411-207-6	Sequence 6, Appli
149	19	100.0	125	12	US-10-424-599-221249	Sequence 221249,
150	19	100.0	125	12	US-10-389-647-495	Sequence 495, App

#### ALIGNMENTS

##### RESULT 1

US-09-169-048-4

; Sequence 4, Application US/09169048A

; Patent No. US20020146740A1

; GENERAL INFORMATION:

; APPLICANT: Huse, William D.

; APPLICANT: Freedman, Michael H.

; TITLE OF INVENTION: Method for Identifying Optimal Binding Ligands to a

; TITLE OF INVENTION: Receptor

; FILE REFERENCE: P-IX 3280

; CURRENT APPLICATION NUMBER: US/09/169,048A

; CURRENT FILING DATE: 1998-10-08

; EARLIER APPLICATION NUMBER: 60/112,011

; EARLIER FILING DATE: 1997-10-09

; NUMBER OF SEQ ID NOS: 28

; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: construct  
US-09-169-048-4

Query Match 100.0%; Score 19; DB 9; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
|||  
Db 3 SVRF 6

RESULT 2

US-09-997-209-4

; Sequence 4, Application US/09997209  
; Publication No. US20030096401A1  
; GENERAL INFORMATION:  
; APPLICANT: Huse, William D.  
; TITLE OF INVENTION: Eukaryotic Expression Libraries and  
; TITLE OF INVENTION: Methods of Use  
; FILE REFERENCE: P-IX 5066  
; CURRENT APPLICATION NUMBER: US/09/997,209  
; CURRENT FILING DATE: 2001-11-28  
; PRIOR APPLICATION NUMBER: US 09/724,762  
; PRIOR FILING DATE: 2000-11-28  
; NUMBER OF SEQ ID NOS: 90  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 8  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic variant  
US-09-997-209-4

Query Match 100.0%; Score 19; DB 10; Length 8;  
Best Local Similarity 100.0%; Pred. No. 9.7e+05;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
|||  
Db 3 SVRF 6

RESULT 3

US-09-839-469-4

; Sequence 4, Application US/09839469  
; Publication No. US20030186316A1  
; GENERAL INFORMATION:  
; APPLICANT: Huse, William D.

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:11:45 ; Search time 5.10924 Seconds  
 (without alignments)  
 75.308 Million cell updates/sec

Title: US-10-030-735-52  
 Perfect score: 19  
 Sequence: 1 SVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	19	100.0	52	2	H85681	unknown protein en
2	19	100.0	55	2	AE2977	hypothetical prote
3	19	100.0	56	2	F85516	unknown protein en
4	19	100.0	66	2	I59652	MHC class II HLA-D
5	19	100.0	72	2	PH0147	HLA-DRB sigma anti
6	19	100.0	73	2	I79665	DRB1 transplantati
7	19	100.0	73	2	AG0609	cold shock-like pr
8	19	100.0	75	2	E95878	hypothetical prote
9	19	100.0	76	2	I68913	MHC protein - cott
10	19	100.0	77	2	C84433	proteinase inhibit
11	19	100.0	78	2	I59643	gene HLA-DRB1 prot
12	19	100.0	78	2	F87158	conserved hypothet
13	19	100.0	78	2	D70751	hypothetical prote

14	19	100.0	80	2	I38020	MHC class II histo
15	19	100.0	84	2	H82860	hypothetical prote
16	19	100.0	87	2	A72210	conserved hypothet
17	19	100.0	89	2	T23831	hypothetical prote
18	19	100.0	89	2	T51192	small zinc finger-
19	19	100.0	91	2	A72597	hypothetical prote
20	19	100.0	92	2	B97809	hypothetical prote
21	19	100.0	93	2	A97190	hypothetical prote
22	19	100.0	95	2	D64725	fixX protein - Esc
23	19	100.0	95	2	G85485	probable ferredoxi
24	19	100.0	95	2	G90634	probable ferredoxi
25	19	100.0	95	2	AH0511	ferredoxin like pr
26	19	100.0	95	2	AF1069	hypothetical prote
27	19	100.0	102	2	S50397	probable membrane
28	19	100.0	104	2	S55062	protein L-aspartyl
29	19	100.0	105	2	JN0864	hypothetical 11.3K
30	19	100.0	107	2	T02814	thioredoxin TRXRP1
31	19	100.0	108	2	F72649	hypothetical prote
32	19	100.0	112	2	AF2627	hypothetical prote
33	19	100.0	112	4	S59290	hypothetical prote
34	19	100.0	114	2	T36142	hypothetical prote
35	19	100.0	115	2	S75854	hypothetical prote
36	19	100.0	115	2	S53390	probable membrane
37	19	100.0	118	2	S69132	Ig heavy chain V-I
38	19	100.0	123	2	A25239	class II histocomp
39	19	100.0	123	2	T21739	hypothetical prote
40	19	100.0	124	2	C69729	urease (beta subun
41	19	100.0	125	2	T25729	hypothetical prote
42	19	100.0	125	2	C83361	hypothetical prote
43	19	100.0	125	2	AF2623	hypothetical prote
44	19	100.0	125	2	E97405	hypothetical prote
45	19	100.0	127	2	H69171	hypothetical prote
46	19	100.0	129	2	S48902	hypothetical prote
47	19	100.0	131	2	G70541	hypothetical prote
48	19	100.0	132	1	QQCVP3	AL3 protein - pota
49	19	100.0	132	2	E82403	hypothetical prote
50	19	100.0	133	2	T38193	very hypothetical
51	19	100.0	133	2	S69568	hypothetical prote
52	19	100.0	134	2	T11385	NADH2 dehydrogenas
53	19	100.0	134	2	T21058	hypothetical prote
54	19	100.0	134	2	B97477	hypothetical prote
55	19	100.0	134	2	AB2695	hypothetical prote
56	19	100.0	137	1	FXME	flavodoxin - Megas
57	19	100.0	140	2	T43309	actin related prot
58	19	100.0	140	2	T50140	probable cytochrom
59	19	100.0	143	2	F82186	hypothetical prote
60	19	100.0	145	2	JC5525	1 alpha,25-dihydro
61	19	100.0	147	2	B75275	hypothetical prote
62	19	100.0	147	2	B27893	hypothetical prote
63	19	100.0	149	2	B69225	hypothetical prote
64	19	100.0	150	2	E72558	hypothetical prote
65	19	100.0	150	2	A70919	hypothetical prote
66	19	100.0	151	1	B39697	myb-related protei
67	19	100.0	152	1	ZDBPF4	gene D protein - p
68	19	100.0	152	2	JS0454	gene D protein - p
69	19	100.0	152	2	D95300	hypothetical prote
70	19	100.0	153	1	OTBY5A	cytochrome-c oxida

71	19	100.0	156	2	S76395	hypothetical prote
72	19	100.0	157	2	T27453	hypothetical prote
73	19	100.0	157	2	C71060	hypothetical prote
74	19	100.0	158	2	S60084	ywiD protein - Bac
75	19	100.0	158	2	T22317	hypothetical prote
76	19	100.0	159	2	F81219	NADH2 dehydrogenas
77	19	100.0	159	2	A56604	TK 5'-region hypot
78	19	100.0	160	2	F72322	hypothetical prote
79	19	100.0	161	2	H84775	hypothetical prote
80	19	100.0	163	2	T47394	hypothetical prote
81	19	100.0	165	2	B81135	E16-related protei
82	19	100.0	166	1	KRBO2B	keratin, 68K type
83	19	100.0	167	2	AE3385	acetyltransferase
84	19	100.0	168	2	T39069	arp 2-3 complex 20
85	19	100.0	169	2	T02943	ubiquitin-conjugat
86	19	100.0	169	2	T01329	ubiquitin-conjugat
87	19	100.0	171	2	T00915	hypothetical prote
88	19	100.0	171	2	B86719	hypothetical prote
89	19	100.0	172	2	A86752	prophage pi2 prote
90	19	100.0	173	2	S70521	cathelin-related p
91	19	100.0	173	2	S76779	hypothetical prote
92	19	100.0	175	2	C90631	hypothetical prote
93	19	100.0	176	2	A45328	bactenecin 5 precu
94	19	100.0	176	2	JT0616	cystatin-related p
95	19	100.0	178	2	S39875	hypothetical prote
96	19	100.0	179	2	E69406	conserved hypothet
97	19	100.0	182	2	T35714	hypothetical prote
98	19	100.0	183	2	T34820	deoxyuridine 5'-tr
99	19	100.0	183	2	T40915	probable RNA-bindi
100	19	100.0	183	2	B75204	probable phosphotr
101	19	100.0	185	2	T34286	hypothetical prote
102	19	100.0	187	2	T18844	hypothetical prote
103	19	100.0	188	1	WZBEI1	dUTP diphosphatase
104	19	100.0	189	2	A43739	development-specif
105	19	100.0	191	2	T31903	hypothetical prote
106	19	100.0	192	2	T05159	hypothetical prote
107	19	100.0	195	2	H95254	hypothetical prote
108	19	100.0	195	2	G98119	hypothetical prote
109	19	100.0	199	2	H90024	hypothetical prote
110	19	100.0	201	2	B72380	hypothetical prote
111	19	100.0	202	2	B83251	glutamine amidotra
112	19	100.0	202	2	B75300	hypothetical prote
113	19	100.0	205	2	T16587	hypothetical prote
114	19	100.0	206	2	T11626	hypothetical prote
115	19	100.0	206	2	G87251	hypothetical prote
116	19	100.0	207	2	C84631	nodulin-like prote
117	19	100.0	207	2	C85482	hypothetical prote
118	19	100.0	207	2	A69941	capsular polysacch
119	19	100.0	210	2	A71187	probable ribosomal
120	19	100.0	210	2	C75147	ssu ribosomal prot
121	19	100.0	210	2	H82037	conserved hypothet
122	19	100.0	210	2	A57337	esaI protein - Erw
123	19	100.0	211	2	F83193	L-isoaspartate pro
124	19	100.0	213	2	A99390	transcriptional ac
125	19	100.0	214	2	B83147	hypothetical prote
126	19	100.0	214	2	S70175	yenI protein - Yer
127	19	100.0	216	2	S35945	EagI protein - Ent

128	19	100.0	216	2	S39625	EagI protein - Ent
129	19	100.0	216	2	AI0120	N-acylhomoserine 1
130	19	100.0	218	2	F64806	ybgJ protein - Esc
131	19	100.0	218	2	H90720	probable carboxyla
132	19	100.0	218	2	F85571	probable carboxyla
133	19	100.0	218	2	AD0588	conserved hypothet
134	19	100.0	218	2	AI0299	N-acylhomoserine 1
135	19	100.0	219	2	B69540	L-isoaspartyl prot
136	19	100.0	219	2	B81534	hypothetical prote
137	19	100.0	221	2	T52138	eukaryotic cap-bin
138	19	100.0	225	2	B72749	probable proteasom
139	19	100.0	226	1	SAVLHV	major surface anti
140	19	100.0	226	1	JQ1577	major surface anti
141	19	100.0	226	1	JQ1581	major surface anti
142	19	100.0	226	1	SAVLN1	major surface anti
143	19	100.0	226	1	SAVLAD	major surface anti
144	19	100.0	226	1	SAVLAR	major surface anti
145	19	100.0	226	1	JQ1570	major surface anti
146	19	100.0	226	2	JQ2045	surface antigen -
147	19	100.0	226	2	JQ2052	surface antigen -
148	19	100.0	226	2	JQ2109	surface antigen -
149	19	100.0	226	2	JQ2114	surface antigen -
150	19	100.0	226	2	JQ2106	surface antigen -

#### ALIGNMENTS

##### RESULT 1

H85681

unknown protein encoded by prophage CP-933N [imported] - Escherichia coli  
(strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C;Accession: H85681

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.  
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: H85681

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-52 <STO>

A;Cross-references: GB:AE005174; NID:g12514732; PIDN:AAG55916.1; GSPDB:GN00145; UWGP:Z1816

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: Z1816

Query Match 100.0%; Score 19; DB 2; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
    ||||  
Db 2 SVRF 5

RESULT 2

AE2977

hypothetical protein Atu3422 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002

C;Accession: AE2977

R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitajima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tingey, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.

A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A;Reference number: AB2577; MUID:21608550; PMID:11743193

A;Accession: AE2977

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-55 <KUR>

A;Cross-references: GB:AE008689; PIDN:AAL44235.1; PID:g17741817; GSPDB:GN00187

A;Experimental source: strain C58 (Dupont)

C;Genetics:

A;Gene: Atu3422

A;Map position: linear chromosome

Query Match 100.0%; Score 19; DB 2; Length 55;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
    ||||  
Db 14 SVRF 17

RESULT 3

F85516

unknown protein encoded in prophage CP-933I [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C;Accession: F85516

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.

Nature 409, 529-533, 2001



OM protein - protein search, using sw model

Run on: April 7, 2004, 18:58:36 ; Search time 2.72269 Seconds  
 (without alignments)  
 76.498 Million cell updates/sec

Title: US-10-030-735-52  
 Perfect score: 19  
 Sequence: 1 SVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	19	100.0	45	1	PSBK_EUGVI	Q9ms55 euglena vir
2	19	100.0	65	1	YH72_PASMU	Q9ck62 pasteurella
3	19	100.0	77	1	THG3_ARATH	Q9zul7 arabidopsis
4	19	100.0	78	1	Y100_MYCTU	Q10895 mycobacteri
5	19	100.0	89	1	IM9A_MOUSE	Q9wv98 mus musculu
6	19	100.0	95	1	FIXX_ECOLI	P31576 escherichia
7	19	100.0	102	1	YM21_YEAST	P40211 saccharomyc
8	19	100.0	113	1	Y544_METTH	O26644 methanobact
9	19	100.0	114	1	YA80_SULTO	Q972q3 sulfolobus
10	19	100.0	117	1	NU3M_SARGL	O63850 sarcophyton
11	19	100.0	118	1	HV3V_HUMAN	P80419 homo sapien
12	19	100.0	120	1	TLB4_MOUSE	P56844 mus musculu
13	19	100.0	121	1	RK14_CYACA	Q9tlu2 cyanidium c
14	19	100.0	124	1	URE2_BACSU	P71035 bacillus su
15	19	100.0	127	1	SECE_BUCAP	Q8ka64 buchnera ap
16	19	100.0	129	1	YHW2_YEAST	P38857 saccharomyc
17	19	100.0	132	1	VAL3_PYMVV	P27264 potato yell

18	19	100.0	137	1	FLAV_MEGEL	P00321	megasphaera
19	19	100.0	140	1	COX6_SCHPO	Q9utf6	schizosacch
20	19	100.0	147	1	YSMA_BACSU	P11469	bacillus su
21	19	100.0	150	1	YH51_AERPE	Q9yb44	aeropyrum p
22	19	100.0	151	1	VGD_BPPHX	P03637	bacterioph
23	19	100.0	153	1	CX5A_YEAST	P00424	saccharomyc
24	19	100.0	156	1	SSB1_STRCO	Q9kyi9	streptomyce
25	19	100.0	158	1	VH21_SFVKA	Q9q909	shope fibro
26	19	100.0	158	1	YWID_BACSU	P46910	bacillus su
27	19	100.0	160	1	RSBW_BACLI	O50231	bacillus li
28	19	100.0	166	1	K2C5_BOVIN	P04262	bos taurus
29	19	100.0	168	1	AR20_SCHPO	Q92352	schizosacch
30	19	100.0	173	1	CRAM_MOUSE	P51437	mus musculu
31	19	100.0	176	1	22P2_RAT	P22283	rattus norv
32	19	100.0	176	1	BCT5_BOVIN	P19660	bos taurus
33	19	100.0	176	1	BCT5_CAPHI	P82018	capra hircu
34	19	100.0	183	1	DUT_STRCO	O54134	streptomyce
35	19	100.0	183	1	KPTA_PYRAB	Q9v2b7	pyrococcus
36	19	100.0	188	1	DUT_HSVI1	P28893	ictalurid h
37	19	100.0	189	1	LVN1_LYTVA	P15262	lytechinus
38	19	100.0	198	1	HB2G_HUMAN	P01911	homo sapien
39	19	100.0	199	1	MOBA_STAAM	Q99s03	staphylococ
40	19	100.0	199	1	MOBA_STAAM	Q8nva4	staphylococ
41	19	100.0	202	1	HI52_PSEAE	P72138	pseudomonas
42	19	100.0	207	1	YPQP_BACSU	P54183	bacillus su
43	19	100.0	210	1	ESAI_ERWST	P54656	erwinia ste
44	19	100.0	210	1	RS3_PYRAB	Q9vlu1	pyrococcus
45	19	100.0	210	1	RS3_PYRHO	O59424	pyrococcus
46	19	100.0	211	1	PIMT_PSEAE	P45683	pseudomonas
47	19	100.0	212	1	ECHI_ERWCH	Q46968	erwinia chr
48	19	100.0	212	1	EXPI_ERWCH	Q47187	erwinia chr
49	19	100.0	213	1	URE1_HELMU	P50044	helicobacte
50	19	100.0	214	1	NADD_PSEAE	Q9hx21	pseudomonas
51	19	100.0	214	1	YENI_YEREN	P52988	yersinia en
52	19	100.0	216	1	EAGI_ENTAG	P33881	enterobacte
53	19	100.0	216	1	YUKI_YERRU	O87970	yersinia ru
54	19	100.0	217	1	COQ7_HUMAN	Q99807	homo sapien
55	19	100.0	218	1	YBGJ_ECOLI	P75744	escherichia
56	19	100.0	219	1	PIM2_ARCFU	O27962	archaeoglob
57	19	100.0	221	1	HIS5_SHEON	Q8efb4	shewanella
58	19	100.0	221	1	IFE3_ARATH	Q9fk59	arabidopsis
59	19	100.0	221	1	SDFL_MOUSE	Q9esp1	mus musculu
60	19	100.0	222	1	NADD_PSESM	Q87vv7	pseudomonas
61	19	100.0	226	1	VMSA_HPBOV	P31873	hepatitis b
62	19	100.0	226	1	VMSA_HPBOV	P31868	hepatitis b
63	19	100.0	226	1	VMSA_HPBOV	P30019	hepatitis b
64	19	100.0	226	1	VMSA_HPBOV	P31869	hepatitis b
65	19	100.0	228	1	SEM5_CAEEL	P29355	caenorhabdi
66	19	100.0	228	1	YDFH_ECOLI	Q8fhd1	escherichia
67	19	100.0	228	1	YDFH_ECOLI	P77577	escherichia
68	19	100.0	229	1	EUTQ_SALTY	Q9zfv5	salmonella
69	19	100.0	230	1	NADD_PSEPK	Q88dl5	pseudomonas
70	19	100.0	239	1	PSA1_ECOLI	P18148	escherichia
71	19	100.0	239	1	PSA2_ECOLI	P17976	escherichia
72	19	100.0	245	1	IFE3_HUMAN	O60573	h eukaryoti
73	19	100.0	245	1	IFE3_MOUSE	Q8bmb3	mus musculu
74	19	100.0	248	1	YDFG_SALTY	P40864	salmonella

75	19	100.0	254	1	RNC_TREPA	O83787	treponema p
76	19	100.0	259	1	SPO7_YEAST	P18410	saccharomyc
77	19	100.0	264	1	MOV_P_CGMVS	P25034	cucumber gr
78	19	100.0	264	1	MOV_P_CGMVW	P19522	cucumber gr
79	19	100.0	264	1	RCII_BPP4	P13059	bacterioph
80	19	100.0	265	1	CEAA_MOUSE	Q61400	mus musculu
81	19	100.0	266	1	2B11_HUMAN	P04229	homo sapien
82	19	100.0	267	1	RL4_SULSO	Q9uxa6	sulfolobus
83	19	100.0	275	1	NARA_RAT	P17982	rattus norv
84	19	100.0	275	1	NARB_RAT	P20974	rattus norv
85	19	100.0	276	1	RNT2_ASPOR	P10281	aspergillus
86	19	100.0	276	1	Y266_ARCFU	O29973	archaeoglob
87	19	100.0	279	1	LE32_CAEL	P36573	caenorhabdi
88	19	100.0	279	1	UP3B_BOVIN	Q864v4	bos taurus
89	19	100.0	283	1	LEG1_HAECO	O44126	haemonchus
90	19	100.0	283	1	SFRA_MOUSE	Q60701	mus musculu
91	19	100.0	284	1	YK62_SULTO	Q96yw6	sulfolobus
92	19	100.0	287	1	NARA_MOUSE	P17981	mus musculu
93	19	100.0	288	1	CHR2_VIBCH	Q9ks61	vibrio chol
94	19	100.0	288	1	GIL2_ENTHI	Q03077	entamoeba h
95	19	100.0	289	1	NARB_MOUSE	O35975	mus musculu
96	19	100.0	294	1	RM45_CAEL	Q95y71	caenorhabdi
97	19	100.0	298	1	YIHV_ECOLI	P32143	escherichia
98	19	100.0	302	1	CYSD_VIBCH	Q9kp19	vibrio chol
99	19	100.0	302	1	CYSD_VIBPA	Q87sy0	vibrio para
100	19	100.0	302	1	CYSD_VIBVU	Q8de72	vibrio vuln
101	19	100.0	304	1	Y687_HAEIN	P71356	haemophilus
102	19	100.0	306	1	SRPI_SYNP7	Q55032	synechococc
103	19	100.0	307	1	MMP1_MYCAV	Q48899	mycobacteri
104	19	100.0	307	1	MMP1_MYCLE	P46841	mycobacteri
105	19	100.0	311	1	SRE1_CAEL	Q09213	caenorhabdi
106	19	100.0	311	1	TUS_YERPE	Q916x9	yersinia pe
107	19	100.0	312	1	NRT1_CHICK	P55806	gallus gall
108	19	100.0	312	1	NRT2_CHICK	P55807	gallus gall
109	19	100.0	313	1	YEIC_ECOLI	P30235	escherichia
110	19	100.0	325	1	YXDK_BACSU	P42422	bacillus su
111	19	100.0	326	1	YIHM_ECOLI	P32134	escherichia
112	19	100.0	330	1	E13C_HORVU	Q02126	hordeum vul
113	19	100.0	331	1	PELB_COLGL	O59939	colletotric
114	19	100.0	340	1	BRA1_HUMAN	Q9gzv7	homo sapien
115	19	100.0	341	1	BRA1_MOUSE	Q9esm3	mus musculu
116	19	100.0	341	1	BRA1_RAT	Q9esm2	rattus norv
117	19	100.0	342	1	HRCA_ACHLA	Q813a0	acholeplasm
118	19	100.0	344	1	DCUP_BRAJA	Q89sk1	bradyrhizob
119	19	100.0	344	1	YD33_MYCTU	Q10644	mycobacteri
120	19	100.0	346	1	DHAS_MYCSM	P41404	mycobacteri
121	19	100.0	346	1	XYLD_RHIME	Q92mt4	rhizobium m
122	19	100.0	350	1	ENG_C_PASMU	Q9cmd1	pasteurella
123	19	100.0	351	1	LPSE_RHIME	Q9r9n2	rhizobium m
124	19	100.0	351	1	MURB_BORPE	Q9x6y8	bordetella
125	19	100.0	357	1	LEU3_SOLTU	P29696	solanum tub
126	19	100.0	360	1	MOA2_MYCTU	O53881	mycobacteri
127	19	100.0	361	1	AROB_LEPIN	Q8exx1	leptospira
128	19	100.0	362	1	FKBH_SCHPO	Q10175	schizosacch
129	19	100.0	362	1	YD33_MYCLE	P53425	mycobacteri
130	19	100.0	365	1	MURG_RALSO	Q8xvi7	ralstonia s
131	19	100.0	366	1	YF59_STAAM	Q99tt7	staphylococ

132	19	100.0	368	1	SNC1_HUMAN	Q16533	homo sapien
133	19	100.0	368	1	YCZ2_YEAST	P25608	saccharomyc
134	19	100.0	370	1	LIPA_SCHPO	O13642	schizosacch
135	19	100.0	371	1	MMK2_MEDSA	Q40353	medicago sa
136	19	100.0	376	1	SSUD_BACSU	P40402	bacillus su
137	19	100.0	376	1	YL60_YEAST	P54007	saccharomyc
138	19	100.0	376	1	YNN4_YEAST	P53912	saccharomyc
139	19	100.0	379	1	ALR_RHIME	Q92r07	rhizobium m
140	19	100.0	382	1	RFC_SHIFL	P37784	shigella fl
141	19	100.0	386	1	KPR3_ARATH	O64888	arabidopsis
142	19	100.0	388	1	ALR3_RHILO	Q981h7	rhizobium l
143	19	100.0	389	1	VMSA_HPBI	P17397	hepatitis b
144	19	100.0	389	1	VMSA_HPBI	P17398	hepatitis b
145	19	100.0	389	1	VMSA_HPBI	P12911	hepatitis b
146	19	100.0	389	1	VMSA_HPBI	P17399	hepatitis b
147	19	100.0	389	1	VMSA_HPBI	P03142	hepatitis b
148	19	100.0	390	1	NCF1_HUMAN	P14598	homo sapien
149	19	100.0	390	1	NCF1_MOUSE	Q09014	mus musculu
150	19	100.0	392	1	NCF1_BOVIN	O77774	bos taurus

#### ALIGNMENTS

##### RESULT 1

##### PSBK\_EUGVI

ID PSBK\_EUGVI STANDARD; PRT; 45 AA.

AC Q9MS55;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Photosystem II reaction center protein K precursor (PSII-K).

GN PSBK.

OS Euglena viridis.

OG Chloroplast.

OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.

OX NCBI\_TaxID=3040;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=21080550; PubMed=11212923;

RA Doetsch N.A., Thompson M.D., Favreau M.R., Hallick R.B.;

RT "Comparison of psbK operon organization and group III intron content

RT in chloroplast genomes of 12 Euglenoid species.";

RL Mol. Gen. Genet. 264:682-690(2001).

CC -!- FUNCTION: This protein is a component of the reaction center of  
CC photosystem II.

CC -!- SIMILARITY: Belongs to the psbK family.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL; AF241284; AAF82462.1; -.

DR HAMAP; MF\_00441; -; 1.  
 DR InterPro; IPR003687; PSII\_PsbK.  
 DR Pfam; PF02533; PsbK; 1.  
 KW Photosystem II; Chloroplast.  
 FT PROPEP 1 8 BY SIMILARITY.  
 FT CHAIN 9 45 PHOTOSYSTEM II REACTION CENTER PROTEIN K.  
 SQ SEQUENCE 45 AA; 5180 MW; 2DC0CD5108224887 CRC64;

Query Match 100.0%; Score 19; DB 1; Length 45;  
 Best Local Similarity 100.0%; Pred. No. 28;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
 ||||  
 Db 41 SVRF 44

# RESULT 2

YH72\_PASMU

ID YH72\_PASMU STANDARD; PRT; 65 AA.

AC Q9CK62;

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein PM1772.

GN PM1772.

OS Pasteurella multocida.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;

OC Pasteurellaceae; Pasteurella.

OX NCBI\_TaxID=747;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Pm70;

RX MEDLINE=21145866; PubMed=11248100;

RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;

RT "Complete genomic sequence of Pasteurella multocida Pm70.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AE006214; AAK03856.1; -.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 65 AA; 7117 MW; F19B147DB954900D CRC64;

Query Match 100.0%; Score 19; DB 1; Length 65;  
 Best Local Similarity 100.0%; Pred. No. 41;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
 ||||  
 Db 48 SVRF 51

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:09:21 ; Search time 16.1345 Seconds  
 (without alignments)  
 78.222 Million cell updates/sec

Title: US-10-030-735-52  
 Perfect score: 19  
 Sequence: 1 SVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertebrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Match	Length	DB	ID	Description
No.						

---

1	19	100.0	34	2	Q9ZG56	Q9zg56 chlamydia t
2	19	100.0	37	16	Q8KCP3	Q8kcp3 chlorobium
3	19	100.0	44	16	Q8F9N3	Q8f9n3 leptospira
4	19	100.0	49	5	Q26786	Q26786 trypanosoma
5	19	100.0	49	16	Q8FGR8	Q8fgr8 escherichia
6	19	100.0	52	10	Q8H3N6	Q8h3n6 oryza sativ
7	19	100.0	52	16	Q8X3R5	Q8x3r5 escherichia
8	19	100.0	53	16	Q8XVC5	Q8xvc5 ralstonia s.
9	19	100.0	53	16	Q8R8G3	Q8r8g3 thermoanaer
10	19	100.0	53	16	Q8CL92	Q8cl92 yersinia pe
11	19	100.0	54	10	Q9LE24	Q9le24 spinacia ol
12	19	100.0	55	16	Q8UAF2	Q8uaf2 agrobacteri
13	19	100.0	55	16	Q8F8Q8	Q8f8q8 leptospira
14	19	100.0	56	16	Q8X3U7	Q8x3u7 escherichia
15	19	100.0	57	10	Q84SN5	Q84sn5 oryza sativ
16	19	100.0	57	12	Q81104	Q81104 hepatitis b
17	19	100.0	58	2	Q50071	Q50071 mycobacteri
18	19	100.0	58	7	Q8WLS0	Q8wls0 homo sapien
19	19	100.0	61	9	Q9B038	Q9b038 mycobacteri
20	19	100.0	61	12	O90266	O90266 hepatitis b
21	19	100.0	61	12	O90264	O90264 hepatitis b
22	19	100.0	61	12	O90265	O90265 hepatitis b
23	19	100.0	61	12	O90260	O90260 hepatitis b
24	19	100.0	61	12	O90268	O90268 hepatitis b
25	19	100.0	61	12	O90259	O90259 hepatitis b
26	19	100.0	61	12	O90263	O90263 hepatitis b
27	19	100.0	61	12	O90262	O90262 hepatitis b
28	19	100.0	61	12	O90267	O90267 hepatitis b
29	19	100.0	61	12	O90258	O90258 hepatitis b
30	19	100.0	61	12	O90261	O90261 hepatitis b
31	19	100.0	61	12	Q80P43	Q80p43 narcissus c
32	19	100.0	61	12	Q80P39	Q80p39 narcissus c
33	19	100.0	62	2	Q7X1L7	Q7x1l7 leptospiril
34	19	100.0	62	16	Q81PB5	Q81pb5 bacillus an
35	19	100.0	63	12	Q9Q5P1	Q9q5p1 hepatitis b
36	19	100.0	63	12	Q9Q5X0	Q9q5x0 hepatitis b
37	19	100.0	63	12	Q9Q678	Q9q678 hepatitis b
38	19	100.0	63	12	Q9Q6C5	Q9q6c5 hepatitis b
39	19	100.0	63	12	Q9Q5Q7	Q9q5q7 hepatitis b
40	19	100.0	63	12	Q9Q6A7	Q9q6a7 hepatitis b
41	19	100.0	63	12	Q9Q5P9	Q9q5p9 hepatitis b
42	19	100.0	63	12	Q9Q5M4	Q9q5m4 hepatitis b
43	19	100.0	63	12	Q9Q613	Q9q613 hepatitis b
44	19	100.0	63	12	Q9Q6D5	Q9q6d5 hepatitis b
45	19	100.0	63	12	Q9Q5Q6	Q9q5q6 hepatitis b
46	19	100.0	63	12	Q9Q660	Q9q660 hepatitis b
47	19	100.0	63	12	Q9Q5S1	Q9q5s1 hepatitis b
48	19	100.0	63	12	Q9Q5Y4	Q9q5y4 hepatitis b
49	19	100.0	63	12	Q9Q637	Q9q637 hepatitis b
50	19	100.0	63	12	Q9Q5V0	Q9q5v0 hepatitis b
51	19	100.0	63	12	Q9Q6A0	Q9q6a0 hepatitis b
52	19	100.0	63	12	Q9Q5W8	Q9q5w8 hepatitis b
53	19	100.0	63	12	Q9Q633	Q9q633 hepatitis b
54	19	100.0	63	12	Q9Q5X9	Q9q5x9 hepatitis b
55	19	100.0	63	12	Q9Q6B5	Q9q6b5 hepatitis b
56	19	100.0	63	12	Q9Q656	Q9q656 hepatitis b
57	19	100.0	63	12	Q9Q5P7	Q9q5p7 hepatitis b

58	19	100.0	63	12	Q9Q5V3	Q9q5v3	hepatitis b
59	19	100.0	63	12	Q9Q5T3	Q9q5t3	hepatitis b
60	19	100.0	63	12	Q9Q692	Q9q692	hepatitis b
61	19	100.0	63	12	Q9Q651	Q9q651	hepatitis b
62	19	100.0	63	12	Q9Q604	Q9q604	hepatitis b
63	19	100.0	63	12	Q9Q6A8	Q9q6a8	hepatitis b
64	19	100.0	63	12	Q9Q5V9	Q9q5v9	hepatitis b
65	19	100.0	63	12	Q9Q5W0	Q9q5w0	hepatitis b
66	19	100.0	63	12	Q9Q5N5	Q9q5n5	hepatitis b
67	19	100.0	63	12	Q9Q5R6	Q9q5r6	hepatitis b
68	19	100.0	63	12	Q9Q5P6	Q9q5p6	hepatitis b
69	19	100.0	63	12	Q9Q664	Q9q664	hepatitis b
70	19	100.0	63	12	Q9Q632	Q9q632	hepatitis b
71	19	100.0	63	12	Q9Q5Y6	Q9q5y6	hepatitis b
72	19	100.0	63	12	Q9Q668	Q9q668	hepatitis b
73	19	100.0	63	12	Q9Q5V2	Q9q5v2	hepatitis b
74	19	100.0	63	12	Q9Q622	Q9q622	hepatitis b
75	19	100.0	63	12	Q9Q5M0	Q9q5m0	hepatitis b
76	19	100.0	63	12	Q9Q630	Q9q630	hepatitis b
77	19	100.0	63	12	Q9Q5R9	Q9q5r9	hepatitis b
78	19	100.0	63	12	Q9Q698	Q9q698	hepatitis b
79	19	100.0	63	12	Q9Q5L9	Q9q5l9	hepatitis b
80	19	100.0	63	12	Q9Q5Q4	Q9q5q4	hepatitis b
81	19	100.0	63	12	Q9Q5W4	Q9q5w4	hepatitis b
82	19	100.0	63	12	Q9Q5R1	Q9q5r1	hepatitis b
83	19	100.0	63	12	Q9Q5Z6	Q9q5z6	hepatitis b
84	19	100.0	63	12	Q9Q686	Q9q686	hepatitis b
85	19	100.0	63	12	Q9Q5X5	Q9q5x5	hepatitis b
86	19	100.0	63	12	Q9Q6D6	Q9q6d6	hepatitis b
87	19	100.0	63	12	Q9Q5X7	Q9q5x7	hepatitis b
88	19	100.0	63	12	Q9Q5Y7	Q9q5y7	hepatitis b
89	19	100.0	63	12	Q9Q611	Q9q611	hepatitis b
90	19	100.0	63	12	Q9Q610	Q9q610	hepatitis b
91	19	100.0	63	12	Q9Q5Q9	Q9q5q9	hepatitis b
92	19	100.0	63	12	Q9Q5T1	Q9q5t1	hepatitis b
93	19	100.0	63	12	Q9Q671	Q9q671	hepatitis b
94	19	100.0	63	12	Q9Q666	Q9q666	hepatitis b
95	19	100.0	63	12	Q9Q5N8	Q9q5n8	hepatitis b
96	19	100.0	63	12	Q9Q5N2	Q9q5n2	hepatitis b
97	19	100.0	63	12	Q9Q601	Q9q601	hepatitis b
98	19	100.0	63	12	Q9Q662	Q9q662	hepatitis b
99	19	100.0	63	12	Q9Q5P2	Q9q5p2	hepatitis b
100	19	100.0	63	12	Q9Q5M2	Q9q5m2	hepatitis b
101	19	100.0	63	12	Q9Q5Q5	Q9q5q5	hepatitis b
102	19	100.0	63	12	Q9Q663	Q9q663	hepatitis b
103	19	100.0	63	12	Q9Q6D1	Q9q6d1	hepatitis b
104	19	100.0	63	12	Q9Q5Y2	Q9q5y2	hepatitis b
105	19	100.0	63	12	Q9Q6B9	Q9q6b9	hepatitis b
106	19	100.0	63	12	Q9Q621	Q9q621	hepatitis b
107	19	100.0	63	12	Q9Q6A4	Q9q6a4	hepatitis b
108	19	100.0	63	12	Q9Q5Z7	Q9q5z7	hepatitis b
109	19	100.0	63	12	Q9Q683	Q9q683	hepatitis b
110	19	100.0	63	12	Q9Q5V5	Q9q5v5	hepatitis b
111	19	100.0	63	12	Q9Q5P8	Q9q5p8	hepatitis b
112	19	100.0	63	12	Q9Q5Q1	Q9q5q1	hepatitis b
113	19	100.0	63	12	Q9Q6A1	Q9q6a1	hepatitis b
114	19	100.0	63	12	Q9Q689	Q9q689	hepatitis b



115	19	100.0	63	12	Q9Q681	Q9q681 hepatitis b
116	19	100.0	63	12	Q9Q5V6	Q9q5v6 hepatitis b
117	19	100.0	63	12	Q9Q659	Q9q659 hepatitis b
118	19	100.0	63	12	Q9Q5Z2	Q9q5z2 hepatitis b
119	19	100.0	63	12	Q9Q619	Q9q619 hepatitis b
120	19	100.0	63	12	Q9Q5R8	Q9q5r8 hepatitis b
121	19	100.0	63	12	Q9Q5M7	Q9q5m7 hepatitis b
122	19	100.0	63	12	Q9Q5L8	Q9q5l8 hepatitis b
123	19	100.0	63	12	Q9Q6B7	Q9q6b7 hepatitis b
124	19	100.0	63	12	Q9Q5S5	Q9q5s5 hepatitis b
125	19	100.0	63	12	Q9Q676	Q9q676 hepatitis b
126	19	100.0	63	12	Q9Q5X1	Q9q5x1 hepatitis b
127	19	100.0	63	12	Q9Q6D8	Q9q6d8 hepatitis b
128	19	100.0	63	12	Q9Q5W3	Q9q5w3 hepatitis b
129	19	100.0	63	12	Q9Q5V4	Q9q5v4 hepatitis b
130	19	100.0	63	12	Q9Q5N4	Q9q5n4 hepatitis b
131	19	100.0	63	12	Q9Q5W2	Q9q5w2 hepatitis b
132	19	100.0	63	12	Q9Q5M9	Q9q5m9 hepatitis b
133	19	100.0	63	12	Q9Q5S8	Q9q5s8 hepatitis b
134	19	100.0	63	12	Q9Q5U2	Q9q5u2 hepatitis b
135	19	100.0	63	12	Q9Q624	Q9q624 hepatitis b
136	19	100.0	63	12	Q9Q5S3	Q9q5s3 hepatitis b
137	19	100.0	63	12	Q9Q608	Q9q608 hepatitis b
138	19	100.0	63	12	Q9Q6C3	Q9q6c3 hepatitis b
139	19	100.0	63	12	Q9Q645	Q9q645 hepatitis b
140	19	100.0	63	12	Q9Q646	Q9q646 hepatitis b
141	19	100.0	63	12	Q9Q672	Q9q672 hepatitis b
142	19	100.0	63	12	Q9Q5R0	Q9q5r0 hepatitis b
143	19	100.0	63	12	Q9Q5N6	Q9q5n6 hepatitis b
144	19	100.0	63	12	Q9Q628	Q9q628 hepatitis b
145	19	100.0	63	12	Q9Q654	Q9q654 hepatitis b
146	19	100.0	63	12	Q9Q642	Q9q642 hepatitis b
147	19	100.0	63	12	Q9Q5T2	Q9q5t2 hepatitis b
148	19	100.0	63	12	Q9Q5N3	Q9q5n3 hepatitis b
149	19	100.0	63	12	Q9Q679	Q9q679 hepatitis b
150	19	100.0	63	12	Q9Q674	Q9q674 hepatitis b

# ALIGNMENTS

## RESULT 1

Q9ZG56

ID Q9ZG56 PRELIMINARY; PRT; 34 AA.

AC Q9ZG56;

DT 01-MAY-1999 (TrEMBLrel. 10, Created)

DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)

DE Preprotein translocase (Fragment).

GN SECY.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI\_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=L2 434B;

RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;

RT "Gene identification of Chlamydia trachomatis by random DNA  
 RT sequencing.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
 DR EMBL; AF087305; AAD04081.1; -.  
 FT NON\_TER 1 1  
 FT NON\_TER 34 34  
 SQ SEQUENCE 34 AA; 3809 MW; 195CEFB04FC36304 CRC64;

Query Match 100.0%; Score 19; DB 2; Length 34;  
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
 ||||  
 Db 5 SVRF 8

## RESULT 2

### Q8KCP3

ID Q8KCP3 PRELIMINARY; PRT; 37 AA.  
 AC Q8KCP3;  
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
 DE Hypothetical protein CT1370.  
 GN CT1370.  
 OS Chlorobium tepidum.  
 OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;  
 OC Chlorobium.  
 OX NCBI\_TaxID=1097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TLS / ATCC 49652 / DSM 12025;  
 RX MEDLINE=22103685; PubMed=12093901;  
 RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,  
 RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,  
 RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,  
 RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,  
 RA Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,  
 RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,  
 RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;  
 RT "The complete genome sequence of Chlorobium tepidum TLS, a  
 RT photosynthetic, anaerobic, green-sulfur bacterium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).  
 DR EMBL; AE012895; AAM72599.1; -.  
 DR TIGR; CT1370; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 37 AA; 4428 MW; FA0877C34165EDB5 CRC64;

Query Match 100.0%; Score 19; DB 16; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SVRF 4  
 ||||  
 Db 4 SVRF 7

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:16:36 ; Search time 6.28571 Seconds  
 (without alignments)  
 32.853 Million cell updates/sec

Title: US-10-030-735-53  
 Perfect score: 20  
 Sequence: 1 QVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	20	100.0	11	1	US-08-464-531-55	Sequence 55, Appl
2	20	100.0	11	2	US-08-461-598-55	Sequence 55, Appl
3	20	100.0	11	3	US-08-322-137-55	Sequence 55, Appl
4	20	100.0	11	3	US-08-582-333A-66	Sequence 66, Appl
5	20	100.0	36	1	US-08-781-020-3	Sequence 3, Appli
6	20	100.0	36	3	US-09-038-935-3	Sequence 3, Appli
7	20	100.0	53	3	US-09-271-608-14	Sequence 14, Appl
8	20	100.0	53	3	US-09-695-950-14	Sequence 14, Appl
9	20	100.0	53	3	US-09-696-147-14	Sequence 14, Appl
10	20	100.0	53	4	US-09-696-364-14	Sequence 14, Appl
11	20	100.0	64	4	US-09-107-532A-4834	Sequence 4834, Ap

12	20	100.0	78	4	US-09-071-035-352	Sequence 352, App
13	20	100.0	105	4	US-09-071-035-350	Sequence 350, App
14	20	100.0	110	4	US-09-134-000C-4583	Sequence 4583, Ap
15	20	100.0	110	4	US-09-134-000C-4821	Sequence 4821, Ap
16	20	100.0	115	4	US-09-107-532A-5066	Sequence 5066, Ap
17	20	100.0	127	4	US-09-134-000C-3539	Sequence 3539, Ap
18	20	100.0	135	4	US-09-328-352-5909	Sequence 5909, Ap
19	20	100.0	136	4	US-09-489-039A-10118	Sequence 10118, A
20	20	100.0	150	4	US-09-732-210-1354	Sequence 1354, Ap
21	20	100.0	150	4	US-09-732-210-1361	Sequence 1361, Ap
22	20	100.0	150	4	US-09-732-210-1365	Sequence 1365, Ap
23	20	100.0	150	4	US-09-732-210-1366	Sequence 1366, Ap
24	20	100.0	150	4	US-09-732-210-1374	Sequence 1374, Ap
25	20	100.0	150	4	US-09-732-210-1378	Sequence 1378, Ap
26	20	100.0	151	4	US-09-732-210-1353	Sequence 1353, Ap
27	20	100.0	162	4	US-09-673-395A-314	Sequence 314, App
28	20	100.0	191	4	US-09-252-991A-26824	Sequence 26824, A
29	20	100.0	196	4	US-09-543-681A-6387	Sequence 6387, Ap
30	20	100.0	197	4	US-09-328-352-4665	Sequence 4665, Ap
31	20	100.0	209	4	US-08-858-207A-386	Sequence 386, App
32	20	100.0	210	4	US-09-252-991A-18398	Sequence 18398, A
33	20	100.0	210	4	US-09-252-991A-31712	Sequence 31712, A
34	20	100.0	218	4	US-09-543-681A-7325	Sequence 7325, Ap
35	20	100.0	219	3	US-08-975-628-2	Sequence 2, Appli
36	20	100.0	219	4	US-09-427-774-2	Sequence 2, Appli
37	20	100.0	232	4	US-09-252-991A-27530	Sequence 27530, A
38	20	100.0	236	4	US-09-252-991A-18160	Sequence 18160, A
39	20	100.0	247	4	US-09-543-681A-5590	Sequence 5590, Ap
40	20	100.0	250	3	US-09-036-987A-9	Sequence 9, Appli
41	20	100.0	250	3	US-09-370-700-9	Sequence 9, Appli
42	20	100.0	250	4	US-09-603-207-9	Sequence 9, Appli
43	20	100.0	253	4	US-09-328-352-8134	Sequence 8134, Ap
44	20	100.0	253	5	PCT-US96-01314-53	Sequence 53, Appl
45	20	100.0	254	4	US-09-252-991A-21339	Sequence 21339, A
46	20	100.0	257	4	US-09-107-532A-3936	Sequence 3936, Ap
47	20	100.0	270	4	US-09-489-039A-11843	Sequence 11843, A
48	20	100.0	275	4	US-09-543-681A-5284	Sequence 5284, Ap
49	20	100.0	280	4	US-09-489-039A-7566	Sequence 7566, Ap
50	20	100.0	285	4	US-09-252-991A-30476	Sequence 30476, A
51	20	100.0	294	4	US-09-252-991A-20997	Sequence 20997, A
52	20	100.0	300	4	US-09-489-039A-13718	Sequence 13718, A
53	20	100.0	301	4	US-09-107-532A-6546	Sequence 6546, Ap
54	20	100.0	301	5	PCT-US95-13975-72	Sequence 72, Appl
55	20	100.0	305	4	US-09-198-452A-883	Sequence 883, App
56	20	100.0	315	4	US-09-489-039A-10235	Sequence 10235, A
57	20	100.0	320	4	US-09-489-039A-11532	Sequence 11532, A
58	20	100.0	332	4	US-09-107-532A-5821	Sequence 5821, Ap
59	20	100.0	334	4	US-09-328-352-4833	Sequence 4833, Ap
60	20	100.0	338	4	US-09-540-236-3196	Sequence 3196, Ap
61	20	100.0	340	4	US-09-252-991A-19963	Sequence 19963, A
62	20	100.0	341	4	US-09-252-991A-29740	Sequence 29740, A
63	20	100.0	342	4	US-09-252-991A-32983	Sequence 32983, A
64	20	100.0	344	4	US-09-328-352-6058	Sequence 6058, Ap
65	20	100.0	346	4	US-09-540-236-2411	Sequence 2411, Ap
66	20	100.0	358	4	US-09-252-991A-32839	Sequence 32839, A
67	20	100.0	359	4	US-09-252-991A-19855	Sequence 19855, A
68	20	100.0	360	4	US-09-252-991A-29934	Sequence 29934, A

69	20	100.0	364	4	US-09-489-039A-8909	Sequence 8909, Ap
70	20	100.0	366	4	US-09-489-039A-10346	Sequence 10346, A
71	20	100.0	367	4	US-09-296-840A-3	Sequence 3, Appli
72	20	100.0	368	4	US-09-252-991A-18904	Sequence 18904, A
73	20	100.0	368	4	US-09-543-681A-8269	Sequence 8269, Ap
74	20	100.0	371	4	US-09-252-991A-23484	Sequence 23484, A
75	20	100.0	371	4	US-09-252-991A-28868	Sequence 28868, A
76	20	100.0	382	1	US-07-783-705A-3	Sequence 3, Appli
77	20	100.0	390	4	US-09-823-823-14	Sequence 14, Appl
78	20	100.0	397	4	US-09-489-039A-12033	Sequence 12033, A
79	20	100.0	403	4	US-09-543-681A-6621	Sequence 6621, Ap
80	20	100.0	404	4	US-09-134-001C-3114	Sequence 3114, Ap
81	20	100.0	407	4	US-09-252-991A-26539	Sequence 26539, A
82	20	100.0	412	4	US-09-252-991A-17613	Sequence 17613, A
83	20	100.0	412	4	US-09-543-681A-8318	Sequence 8318, Ap
84	20	100.0	416	4	US-09-430-221-13	Sequence 13, Appl
85	20	100.0	427	4	US-09-550-645-2	Sequence 2, Appli
86	20	100.0	431	4	US-09-252-991A-33119	Sequence 33119, A
87	20	100.0	435	3	US-08-911-321-8	Sequence 8, Appli
88	20	100.0	442	4	US-09-252-991A-28733	Sequence 28733, A
89	20	100.0	450	4	US-09-489-039A-13998	Sequence 13998, A
90	20	100.0	452	4	US-09-252-991A-25818	Sequence 25818, A
91	20	100.0	454	4	US-09-489-039A-12506	Sequence 12506, A
92	20	100.0	463	2	US-08-846-762-18	Sequence 18, Appl
93	20	100.0	472	4	US-09-252-991A-30367	Sequence 30367, A
94	20	100.0	477	4	US-09-328-352-6549	Sequence 6549, Ap
95	20	100.0	485	4	US-09-172-952-27	Sequence 27, Appl
96	20	100.0	489	4	US-09-252-991A-18918	Sequence 18918, A
97	20	100.0	492	1	US-10-095-946-2	Sequence 2, Appli
98	20	100.0	492	4	US-09-183-959-2	Sequence 2, Appli
99	20	100.0	492	4	US-09-347-650-14	Sequence 14, Appl
100	20	100.0	492	4	US-09-535-315-2	Sequence 2, Appli
101	20	100.0	493	4	US-09-252-991A-31287	Sequence 31287, A
102	20	100.0	497	4	US-09-252-991A-18027	Sequence 18027, A
103	20	100.0	500	4	US-09-328-352-8183	Sequence 8183, Ap
104	20	100.0	502	4	US-09-328-857A-4	Sequence 4, Appli
105	20	100.0	503	4	US-09-252-991A-31626	Sequence 31626, A
106	20	100.0	505	3	US-09-257-581-5	Sequence 5, Appli
107	20	100.0	505	3	US-09-257-581-7	Sequence 7, Appli
108	20	100.0	510	4	US-09-252-991A-25258	Sequence 25258, A
109	20	100.0	522	3	US-09-165-042-3	Sequence 3, Appli
110	20	100.0	533	4	US-09-252-991A-20633	Sequence 20633, A
111	20	100.0	537	4	US-09-489-039A-10024	Sequence 10024, A
112	20	100.0	538	4	US-09-198-452A-144	Sequence 144, App
113	20	100.0	551	4	US-08-952-061-2	Sequence 2, Appli
114	20	100.0	551	4	US-09-252-991A-20358	Sequence 20358, A
115	20	100.0	556	4	US-09-252-991A-18923	Sequence 18923, A
116	20	100.0	563	4	US-09-252-991A-26415	Sequence 26415, A
117	20	100.0	565	4	US-09-252-991A-20122	Sequence 20122, A
118	20	100.0	570	4	US-09-252-991A-16680	Sequence 16680, A
119	20	100.0	573	4	US-09-252-991A-31661	Sequence 31661, A
120	20	100.0	573	4	US-09-489-039A-11884	Sequence 11884, A
121	20	100.0	591	4	US-09-050-739-72	Sequence 72, Appl
122	20	100.0	595	4	US-09-252-991A-17469	Sequence 17469, A
123	20	100.0	602	4	US-09-540-236-3263	Sequence 3263, Ap
124	20	100.0	607	4	US-09-252-991A-30878	Sequence 30878, A
125	20	100.0	608	4	US-09-107-532A-5702	Sequence 5702, Ap

126	20	100.0	616	3	US-09-136-574A-47	Sequence 47, Appl
127	20	100.0	616	4	US-09-489-039A-9231	Sequence 9231, Ap
128	20	100.0	621	2	US-08-419-652-4	Sequence 4, Appli
129	20	100.0	646	4	US-09-252-991A-16819	Sequence 16819, A
130	20	100.0	649	3	US-08-996-441B-109	Sequence 109, App
131	20	100.0	649	3	US-08-993-722A-109	Sequence 109, App
132	20	100.0	649	3	US-08-993-170A-109	Sequence 109, App
133	20	100.0	649	3	US-08-993-775B-109	Sequence 109, App
134	20	100.0	649	4	US-09-252-991A-27657	Sequence 27657, A
135	20	100.0	649	4	US-09-427-770-109	Sequence 109, App
136	20	100.0	649	4	US-09-427-769-109	Sequence 109, App
137	20	100.0	655	4	US-09-311-626B-14	Sequence 14, Appl
138	20	100.0	656	4	US-09-543-681A-6473	Sequence 6473, Ap
139	20	100.0	660	1	US-08-248-532-3	Sequence 3, Appli
140	20	100.0	660	2	US-08-419-652-3	Sequence 3, Appli
141	20	100.0	662	1	US-08-248-532-2	Sequence 2, Appli
142	20	100.0	662	2	US-08-419-652-2	Sequence 2, Appli
143	20	100.0	662	2	US-08-685-118-4	Sequence 4, Appli
144	20	100.0	662	2	US-08-915-495-4	Sequence 4, Appli
145	20	100.0	662	2	US-08-914-520-4	Sequence 4, Appli
146	20	100.0	662	3	US-08-789-350-2	Sequence 2, Appli
147	20	100.0	689	4	US-09-425-335-2	Sequence 2, Appli
148	20	100.0	705	4	US-09-252-991A-23298	Sequence 23298, A
149	20	100.0	719	4	US-09-252-991A-28768	Sequence 28768, A
150	20	100.0	724	4	US-09-543-681A-4745	Sequence 4745, Ap

#### ALIGNMENTS

#### RESULT 1

US-08-464-531-55

; Sequence 55, Application US/08464531

; Patent No. 5789184

; GENERAL INFORMATION:

; APPLICANT: FOWLKES, Dana M.

; APPLICANT: BROACH, Jim

; APPLICANT: MANFREDI, John

; APPLICANT: KLEIN, Christine

; APPLICANT: MURPHY, Andrew J.

; APPLICANT: PAUL, Jeremy

; APPLICANT: TRUEHEART, Joshua

; TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE

; TITLE OF INVENTION: PHERMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR

; NUMBER OF SEQUENCES: 119

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: BROWDY AND NEIMARK

; STREET: 419 Seventh Street, N.W., Suite 300

; CITY: Washington

; STATE: D.C.

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,531
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/322,137
; FILING DATE: 13-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/309,313
; FILING DATE: 20-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,328
; FILING DATE: 31-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,431
; FILING DATE: 31-MAR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: COOPER, Iver P.
; REGISTRATION NUMBER: 28,005
; REFERENCE/DOCKET NUMBER: FOLWKES=2G
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-531-55

```

```

Query Match          100.0%; Score 20; DB 1; Length 11;
Best Local Similarity 100.0%; Pred. No. 24;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QVRF 4
        ||||
Db      4 QVRF 7

```

## RESULT 2

US-08-461-598-55

```

; Sequence 55, Application US/08461598
; Patent No. 5876951

```

### GENERAL INFORMATION:

```

; APPLICANT: FOWLKES, Dana M.
; APPLICANT: BROACH, Jim
; APPLICANT: MANFREDI, John
; APPLICANT: KLEIN, Christine
; APPLICANT: MURPHY, Andrew J.
; APPLICANT: PAUL, Jeremy
; APPLICANT: TRUEHEART, Joshua
; TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
; TITLE OF INVENTION: PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 22:05:26 ; Search time 16.8403 Seconds  
 (without alignments)  
 62.382 Million cell updates/sec

Title: US-10-030-735-53  
 Perfect score: 20  
 Sequence: 1 QVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	%	Query				
No.	Score	Match Length	DB	ID		Description



1	20	100.0	7	14	US-10-283-423-114	Sequence 114, App
2	20	100.0	7	14	US-10-213-821-114	Sequence 114, App
3	20	100.0	11	10	US-09-309-196-55	Sequence 55, Appl
4	20	100.0	11	10	US-09-201-396-36	Sequence 36, Appl
5	20	100.0	11	10	US-09-953-354-66	Sequence 66, Appl
6	20	100.0	11	12	US-09-747-774A-66	Sequence 66, Appl
7	20	100.0	11	14	US-10-277-607-36	Sequence 36, Appl
8	20	100.0	27	14	US-10-295-809-26	Sequence 26, Appl
9	20	100.0	33	9	US-09-864-761-47628	Sequence 47628, A
10	20	100.0	49	12	US-10-424-599-226349	Sequence 226349,
11	20	100.0	52	12	US-10-424-599-219935	Sequence 219935,
12	20	100.0	53	12	US-10-424-599-273990	Sequence 273990,
13	20	100.0	53	13	US-10-043-452-14	Sequence 14, Appl
14	20	100.0	57	11	US-09-864-408A-4524	Sequence 4524, Ap
15	20	100.0	58	9	US-09-864-761-39103	Sequence 39103, A
16	20	100.0	58	12	US-10-424-599-165993	Sequence 165993,
17	20	100.0	59	12	US-10-424-599-148639	Sequence 148639,
18	20	100.0	69	12	US-10-424-599-181866	Sequence 181866,
19	20	100.0	71	12	US-10-424-599-282459	Sequence 282459,
20	20	100.0	72	12	US-10-424-599-143878	Sequence 143878,
21	20	100.0	72	12	US-10-425-114-54359	Sequence 54359, A
22	20	100.0	73	12	US-10-424-599-244507	Sequence 244507,
23	20	100.0	73	12	US-10-424-599-249013	Sequence 249013,
24	20	100.0	76	12	US-10-424-599-235516	Sequence 235516,
25	20	100.0	76	12	US-10-424-599-284498	Sequence 284498,
26	20	100.0	77	12	US-10-424-599-226180	Sequence 226180,
27	20	100.0	81	12	US-10-424-599-279896	Sequence 279896,
28	20	100.0	82	11	US-09-864-408A-8248	Sequence 8248, Ap
29	20	100.0	83	12	US-10-424-599-235865	Sequence 235865,
30	20	100.0	86	12	US-10-424-599-244699	Sequence 244699,
31	20	100.0	90	12	US-10-425-114-42635	Sequence 42635, A
32	20	100.0	91	12	US-10-424-599-159092	Sequence 159092,
33	20	100.0	92	9	US-09-925-300-1793	Sequence 1793, Ap
34	20	100.0	93	12	US-10-424-599-156861	Sequence 156861,
35	20	100.0	102	14	US-10-156-761-7879	Sequence 7879, Ap
36	20	100.0	104	12	US-10-424-599-160320	Sequence 160320,
37	20	100.0	104	12	US-10-425-114-45357	Sequence 45357, A
38	20	100.0	109	12	US-10-424-599-156111	Sequence 156111,
39	20	100.0	111	11	US-09-864-408A-4932	Sequence 4932, Ap
40	20	100.0	114	12	US-10-424-599-242090	Sequence 242090,
41	20	100.0	116	12	US-10-424-599-186719	Sequence 186719,
42	20	100.0	117	12	US-10-424-599-212346	Sequence 212346,
43	20	100.0	120	12	US-10-425-114-39420	Sequence 39420, A
44	20	100.0	120	15	US-10-104-047-3300	Sequence 3300, Ap
45	20	100.0	122	12	US-10-424-599-159298	Sequence 159298,
46	20	100.0	124	12	US-10-424-599-179266	Sequence 179266,
47	20	100.0	127	12	US-10-425-114-37874	Sequence 37874, A
48	20	100.0	127	12	US-10-425-114-58730	Sequence 58730, A
49	20	100.0	128	12	US-10-424-599-150843	Sequence 150843,
50	20	100.0	134	9	US-09-164-615-8	Sequence 8, Appli
51	20	100.0	134	15	US-10-465-046-8	Sequence 8, Appli
52	20	100.0	138	12	US-10-424-599-261163	Sequence 261163,
53	20	100.0	139	15	US-10-108-260A-3323	Sequence 3323, Ap
54	20	100.0	140	12	US-10-424-599-263635	Sequence 263635,
55	20	100.0	143	11	US-09-864-408A-3264	Sequence 3264, Ap
56	20	100.0	143	12	US-10-424-599-243800	Sequence 243800,

57	20	100.0	146	12	US-10-335-977-6050	Sequence 6050, Ap
58	20	100.0	151	12	US-10-221-278-190	Sequence 190, App
59	20	100.0	151	12	US-10-425-114-60310	Sequence 60310, A
60	20	100.0	151	15	US-10-291-172-190	Sequence 190, App
61	20	100.0	153	9	US-09-925-301-1581	Sequence 1581, Ap
62	20	100.0	153	12	US-10-424-599-188065	Sequence 188065,
63	20	100.0	154	12	US-10-424-599-162776	Sequence 162776,
64	20	100.0	157	12	US-10-424-599-215612	Sequence 215612,
65	20	100.0	157	12	US-10-425-114-37449	Sequence 37449, A
66	20	100.0	158	12	US-10-424-599-215614	Sequence 215614,
67	20	100.0	161	12	US-10-425-114-44896	Sequence 44896, A
68	20	100.0	161	12	US-10-425-114-60868	Sequence 60868, A
69	20	100.0	161	12	US-10-425-114-62621	Sequence 62621, A
70	20	100.0	161	12	US-10-425-114-69936	Sequence 69936, A
71	20	100.0	163	12	US-10-424-599-263138	Sequence 263138,
72	20	100.0	164	12	US-10-425-114-46389	Sequence 46389, A
73	20	100.0	166	9	US-09-764-868-1222	Sequence 1222, Ap
74	20	100.0	166	12	US-10-424-599-267551	Sequence 267551,
75	20	100.0	166	12	US-10-425-114-41147	Sequence 41147, A
76	20	100.0	167	9	US-09-734-017A-38	Sequence 38, Appl
77	20	100.0	167	14	US-10-080-170-538	Sequence 538, App
78	20	100.0	172	12	US-10-424-599-230804	Sequence 230804,
79	20	100.0	173	9	US-09-912-020-309	Sequence 309, App
80	20	100.0	173	12	US-10-282-122A-42638	Sequence 42638, A
81	20	100.0	173	12	US-10-282-122A-56307	Sequence 56307, A
82	20	100.0	175	14	US-10-078-770-102	Sequence 102, App
83	20	100.0	175	15	US-10-369-493-9639	Sequence 9639, Ap
84	20	100.0	176	14	US-10-127-032-122	Sequence 122, App
85	20	100.0	181	12	US-10-425-114-38055	Sequence 38055, A
86	20	100.0	182	12	US-10-424-599-202607	Sequence 202607,
87	20	100.0	182	12	US-10-424-599-233700	Sequence 233700,
88	20	100.0	182	12	US-09-925-298-529	Sequence 529, App
89	20	100.0	182	14	US-10-102-806-529	Sequence 529, App
90	20	100.0	183	12	US-10-425-114-43562	Sequence 43562, A
91	20	100.0	186	12	US-10-425-114-64707	Sequence 64707, A
92	20	100.0	186	15	US-10-369-493-11542	Sequence 11542, A
93	20	100.0	186	15	US-10-369-493-14594	Sequence 14594, A
94	20	100.0	186	15	US-10-369-493-14964	Sequence 14964, A
95	20	100.0	186	16	US-10-389-566-1396	Sequence 1396, Ap
96	20	100.0	191	9	US-09-738-973-393	Sequence 393, App
97	20	100.0	191	9	US-09-854-133-393	Sequence 393, App
98	20	100.0	191	14	US-10-144-649A-393	Sequence 393, App
99	20	100.0	192	12	US-10-425-114-38912	Sequence 38912, A
100	20	100.0	193	12	US-10-425-114-37310	Sequence 37310, A
101	20	100.0	194	9	US-09-764-868-676	Sequence 676, App
102	20	100.0	194	9	US-09-764-868-1099	Sequence 1099, Ap
103	20	100.0	194	9	US-09-738-626-4966	Sequence 4966, Ap
104	20	100.0	196	9	US-09-815-242-11660	Sequence 11660, A
105	20	100.0	196	12	US-10-424-599-281558	Sequence 281558,
106	20	100.0	196	12	US-10-425-114-69917	Sequence 69917, A
107	20	100.0	197	9	US-09-738-626-3562	Sequence 3562, Ap
108	20	100.0	197	15	US-10-369-493-11850	Sequence 11850, A
109	20	100.0	200	12	US-10-282-122A-72332	Sequence 72332, A
110	20	100.0	204	14	US-10-180-927-2	Sequence 2, Appli
111	20	100.0	204	14	US-10-156-761-9250	Sequence 9250, Ap
112	20	100.0	205	12	US-10-425-114-39504	Sequence 39504, A
113	20	100.0	210	12	US-10-425-114-50468	Sequence 50468, A

114	20	100.0	210	12	US-10-425-114-59850	Sequence 59850, A
115	20	100.0	214	9	US-09-925-300-1287	Sequence 1287, Ap
116	20	100.0	216	12	US-10-425-114-49301	Sequence 49301, A
117	20	100.0	221	12	US-10-424-599-195388	Sequence 195388, A
118	20	100.0	227	12	US-10-425-114-41216	Sequence 41216, A
119	20	100.0	227	12	US-10-425-114-56779	Sequence 56779, A
120	20	100.0	227	12	US-10-425-114-57022	Sequence 57022, A
121	20	100.0	231	9	US-09-815-242-11998	Sequence 11998, A
122	20	100.0	231	12	US-10-282-122A-66626	Sequence 66626, A
123	20	100.0	231	12	US-10-282-122A-68296	Sequence 68296, A
124	20	100.0	234	15	US-10-369-493-4088	Sequence 4088, Ap
125	20	100.0	239	12	US-10-425-114-39359	Sequence 39359, A
126	20	100.0	239	12	US-10-425-114-65240	Sequence 65240, A
127	20	100.0	240	14	US-10-156-761-13292	Sequence 13292, A
128	20	100.0	243	12	US-10-282-122A-56063	Sequence 56063, A
129	20	100.0	243	12	US-10-282-122A-59653	Sequence 59653, A
130	20	100.0	243	12	US-10-282-122A-69000	Sequence 69000, A
131	20	100.0	243	12	US-10-282-122A-76312	Sequence 76312, A
132	20	100.0	243	12	US-10-282-122A-77053	Sequence 77053, A
133	20	100.0	243	12	US-10-282-122A-77766	Sequence 77766, A
134	20	100.0	243	15	US-10-369-493-19407	Sequence 19407, A
135	20	100.0	250	10	US-09-769-734-2	Sequence 2, Appli
136	20	100.0	250	12	US-10-424-599-220762	Sequence 220762, A
137	20	100.0	250	16	US-10-329-148A-9	Sequence 9, Appli
138	20	100.0	252	9	US-09-815-242-10296	Sequence 10296, A
139	20	100.0	252	12	US-10-282-122A-56679	Sequence 56679, A
140	20	100.0	255	12	US-10-282-122A-73177	Sequence 73177, A
141	20	100.0	256	12	US-10-424-599-240176	Sequence 240176, A
142	20	100.0	259	9	US-09-815-242-13738	Sequence 13738, A
143	20	100.0	259	12	US-10-282-122A-74918	Sequence 74918, A
144	20	100.0	266	12	US-10-425-114-39518	Sequence 39518, A
145	20	100.0	266	14	US-10-156-761-15054	Sequence 15054, A
146	20	100.0	269	12	US-10-282-122A-67133	Sequence 67133, A
147	20	100.0	274	12	US-10-282-122A-57529	Sequence 57529, A
148	20	100.0	276	12	US-10-425-114-38383	Sequence 38383, A
149	20	100.0	278	12	US-10-425-114-66403	Sequence 66403, A
150	20	100.0	280	12	US-10-282-122A-47141	Sequence 47141, A

#### ALIGNMENTS

##### RESULT 1

US-10-283-423-114

; Sequence 114, Application US/10283423

; Publication No. US20030162223A1

; GENERAL INFORMATION:

; APPLICANT: Lowery, David E.

; APPLICANT: Smith, Valdin G.

; APPLICANT: Kubiak, Teresa M.

; APPLICANT: Larsen, Martha J.

; TITLE OF INVENTION: Drosophila G Protein Coupled Receptors, Nucleic Acids, And Methods

; TITLE OF INVENTION: Related To The Same

; FILE REFERENCE: PHRM0002-102

; Application Project

; -----

```

; CURRENT APPLICATION NUMBER: US/10/283,423
; CURRENT FILING DATE: 2002-10-30
;   Earlier Applications
;   -----
; PRIOR APPLICATION NUMBER: PriorAppNumber : 09/693,746
; PRIOR FILING DATE: PriorFilingDate : 2000-10-20
; PRIOR APPLICATION NUMBER: PriorAppNumber : 09/425,676
; PRIOR FILING DATE: PriorFilingDate : 1999-10-22
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 114
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20030162223A1el Sequence
US-10-283-423-114

```

```

Query Match          100.0%; Score 20; DB 14; Length 7;
Best Local Similarity 100.0%; Pred. No. 9.7e+05;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 QVRF 4
        ||||
Db      4 QVRF 7

```

# RESULT 2

```

US-10-213-821-114
; Sequence 114, Application US/10213821
; Publication No. US20030180297A1
; GENERAL INFORMATION:
; APPLICANT: Pharmacia & Upjohn
; APPLICANT: Lowery, David E.
; APPLICANT: Smith, Valdin G.
; APPLICANT: Kubiak, Teresa M.
; APPLICANT: Larsen, Martha J.
; TITLE OF INVENTION: Drosophila G Protein Coupled Receptors, Nucleic Acids,
And Methods
; TITLE OF INVENTION: Related To The Same
; FILE REFERENCE: 6297.2cp
; CURRENT APPLICATION NUMBER: US/10/213,821
; CURRENT FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 09/693,746
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/425,676
; PRIOR FILING DATE: 1999-10-22
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 114
;   LENGTH: 7
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: No. US20030180297A1el Sequence
US-10-213-821-114

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:58:36 ; Search time 2.72269 Seconds  
 (without alignments)  
 76.498 Million cell updates/sec

Title: US-10-030-735-53  
 Perfect score: 20  
 Sequence: 1 QVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	%		DB	ID	Description
		Query	Match Length			
1	20	100.0	15	1	ALLS_MANSE	P42559 manduca sex
2	20	100.0	63	1	YM26_MARPO	P38468 marchantia
3	20	100.0	81	1	Y65_BPT3	P20329 bacteriopha
4	20	100.0	84	1	Y65_BPT7	P03800 bacteriopha
5	20	100.0	102	1	Y66A_HAEIN	O86228 haemophilus
6	20	100.0	114	1	RS13_MUSDO	P27072 musca domes
7	20	100.0	119	1	SMPA_VIBCH	P52117 vibrio chol
8	20	100.0	134	1	VAL3_CLVK	P14977 cassava lat
9	20	100.0	134	1	VAL3_CLVN	P14969 cassava lat
10	20	100.0	140	1	COX6_SCHPO	Q9utf6 schizosacch
11	20	100.0	144	1	PSB1_ECOLI	P10031 escherichia
12	20	100.0	144	1	PSB2_ECOLI	P10032 escherichia
13	20	100.0	150	1	RS13_ANOGA	P52811 anopheles g
14	20	100.0	150	1	RS13_DROME	Q03334 drosophila
15	20	100.0	150	1	RS13_HUMAN	Q02546 homo sapien
16	20	100.0	150	1	RS13 ICTPU	P47772 ictalurus p
17	20	100.0	150	1	RS13_SCHPO	P28189 schizosacch

18	20	100.0	150	1	RS13_XENLA	P49393	xenopus lae
19	20	100.0	151	1	RS13_AGABI	P78571	agaricus bi
20	20	100.0	153	1	RECX_VIBVU	Q8dc50	vibrio vuln
21	20	100.0	155	1	RECX_VIBPA	Q87lr2	vibrio para
22	20	100.0	159	1	RECX_RALSO	Q8yly5	ralstonia s
23	20	100.0	173	1	INSJ_ECOLI	P19768	escherichia
24	20	100.0	173	1	INSJ_SHIFL	P60046	shigella fl
25	20	100.0	175	1	OMLA_PSEFL	O68564	pseudomonas
26	20	100.0	176	1	OMLA_PSEAE	O68562	pseudomonas
27	20	100.0	184	1	BGB_DROME	Q24040	drosophila
28	20	100.0	190	1	GCH1_CAMJE	P51594	campylobact
29	20	100.0	197	1	GCH1_RHILO	Q98lq6	rhizobium l
30	20	100.0	204	1	GCH1_RHIME	Q92qb4	rhizobium m
31	20	100.0	205	1	ALKH_ZYMMO	Q00384	z khg/kdp
32	20	100.0	207	1	YL96_RALSO	P58634	ralstonia s
33	20	100.0	208	1	GCH1_AGR5	Q8uek8	agrobacteri
34	20	100.0	212	1	HIS5_PSEPK	Q88r44	pseudomonas
35	20	100.0	213	1	ALKH_ECOLI	P10177	e khg/kdp
36	20	100.0	213	1	GCH1_BRUME	Q8yh94	brucella me
37	20	100.0	213	1	GCH1_BRUSU	Q8g0l4	brucella su
38	20	100.0	213	1	HI51_PSEAE	Q9hu42	pseudomonas
39	20	100.0	215	1	YAA4_YEAST	P39703	saccharomyc
40	20	100.0	224	1	YHCA_ECOLI	P28722	escherichia
41	20	100.0	229	1	UCK_CABEL	Q17413	caenorhabdi
42	20	100.0	231	1	RL1_PSEAE	Q9hwc6	pseudomonas
43	20	100.0	239	1	LPXH_PASMU	Q9cpe3	pasteurella
44	20	100.0	243	1	YGGJ_ECOLI	P37912	escherichia
45	20	100.0	246	1	VATD_MANSE	Q9u0s4	manduca sex
46	20	100.0	266	1	YAFD_ECOLI	P30865	escherichia
47	20	100.0	266	1	YAFD_SALTI	Q8z985	salmonella
48	20	100.0	266	1	YAFD_SALTY	Q8zrm4	salmonella
49	20	100.0	267	1	PPNK_LACPL	Q88v61	lactobacill
50	20	100.0	269	1	MURI_PASMU	P57887	pasteurella
51	20	100.0	269	1	PYR5_FREDI	P11401	fremyella d
52	20	100.0	270	1	HMPH_HUMAN	Q03014	homo sapien
53	20	100.0	271	1	HMPH_MOUSE	P43120	mus musculu
54	20	100.0	271	1	PYR6_FREDI	P14880	fremyella d
55	20	100.0	275	1	BLOI_PSEAE	O07293	pseudomonas
56	20	100.0	277	1	HMPH_CHICK	Q05502	gallus gall
57	20	100.0	278	1	PYR2_MASLA	P11399	mastigoclad
58	20	100.0	285	1	YD98_YEAST	Q03835	saccharomyc
59	20	100.0	296	1	RL5_MOUSE	P47962	mus musculu
60	20	100.0	296	1	RL5_RAT	P09895	rattus norv
61	20	100.0	325	1	Y272_SYNY3	P73893	synechocyst
62	20	100.0	327	1	RL5_ANOGA	O44248	anopheles g
63	20	100.0	327	1	SYFA_PASMU	P57860	pasteurella
64	20	100.0	328	1	Y002_NPVAC	P24655	autographa
65	20	100.0	329	1	SRA8_CAEEL	Q09210	caenorhabdi
66	20	100.0	329	1	SYFA_HAEIN	P43819	haemophilus
67	20	100.0	330	1	SYFA_NEIMA	Q9j76	neisseria m
68	20	100.0	330	1	SYFA_NEIMB	Q9k092	neisseria m
69	20	100.0	333	1	RRS1_CAEEL	Q9xvt0	caenorhabdi
70	20	100.0	336	1	YERI_BACSU	O34640	bacillus su
71	20	100.0	339	1	TCMO_STRGA	P39896	streptomyce
72	20	100.0	340	1	MOAA_STACA	Q9zim6	staphylococ
73	20	100.0	340	1	REP4_ECOLI	P13961	escherichia
74	20	100.0	343	1	REP7_ECOLI	Q52221	escherichia

75	20	100.0	343	1	REP8_ECOLI	Q52346	escherichia
76	20	100.0	343	1	REPZ_ECOLI	P18023	escherichia
77	20	100.0	344	1	SYFA_RALSO	Q8xz25	ralstonia s
78	20	100.0	353	1	MURG_HELPJ	Q9zk59	helicobacte
79	20	100.0	353	1	MURG_HELPY	O25770	helicobacte
80	20	100.0	353	1	PMA1_MOUSE	Q8c1c8	mus musculu
81	20	100.0	353	1	PMA1_RAT	Q8vhz4	rattus norv
82	20	100.0	365	1	FLGI_ECO57	P58203	escherichia
83	20	100.0	365	1	FLGI_ECOLI	P75941	escherichia
84	20	100.0	365	1	FLGI_SALTY	P15930	salmonella
85	20	100.0	367	1	DP3B_PSEAE	Q9i7c4	pseudomonas
86	20	100.0	367	1	FLGI_SALTI	Q8z7k2	salmonella
87	20	100.0	367	1	MTC1_CHVI3	P36216	chlorella v
88	20	100.0	369	1	FLGI_YERPE	Q8zfb1	yersinia pe
89	20	100.0	373	1	NSDL_HUMAN	Q15738	homo sapien
90	20	100.0	378	1	TRMU_XANAC	Q8pl08	xanthomonas
91	20	100.0	381	1	SSUD_PSEAE	Q9hyg2	pseudomonas
92	20	100.0	382	1	CRTY_PANAN	P21687	pantoea ana
93	20	100.0	390	1	ASSY_SULTO	Q970v0	sulfolobus
94	20	100.0	395	1	ASSY_RHOFA	Q93jq8	rhodococcus
95	20	100.0	396	1	ASSY_STRA3	Q8e7n1	streptococc
96	20	100.0	396	1	ASSY_STRA5	Q8e272	streptococc
97	20	100.0	396	1	ASSY_STRMU	Q8cwz0	streptococc
98	20	100.0	396	1	RF1M_KLULA	P41767	kluyveromyc
99	20	100.0	397	1	ASSY_PYRAE	Q8zu97	pyrobaculum
100	20	100.0	397	1	ASSY_STRCL	P50986	streptomyce
101	20	100.0	397	1	ASSY_THEVO	Q97a55	thermoplasm
102	20	100.0	397	1	ASSY_XANCP	Q8p8j4	xanthomonas
103	20	100.0	398	1	ASSY_CLOTE	P59602	clostridium
104	20	100.0	398	1	ASSY_LACLA	P57799	lactococcus
105	20	100.0	398	1	ASSY_MYCTU	P94993	mycobacteri
106	20	100.0	398	1	ASSY_STRR6	Q8dri5	streptococc
107	20	100.0	399	1	ASSY_MYCLE	Q9cc10	mycobacteri
108	20	100.0	400	1	ASSY_ANASP	Q8ymx6	anabaena sp
109	20	100.0	400	1	ASSY_CLOAB	Q97ke6	clostridium
110	20	100.0	400	1	ASSY_SYNY3	P77973	synechocyst
111	20	100.0	400	1	ASSY_THETH	P59846	thermus the
112	20	100.0	401	1	ASSY_AQUAE	O67213	aquifex aeo
113	20	100.0	401	1	ASSY_BACAA	Q81kv7	bacillus an
114	20	100.0	401	1	ASSY_BACCR	Q817c6	bacillus ce
115	20	100.0	401	1	ASSY_BUCAP	Q8ka60	buchnera ap
116	20	100.0	401	1	ASSY_CHLTE	Q8kde0	chlorobium
117	20	100.0	401	1	ASSY_COREF	Q8ftm9	corynebacte
118	20	100.0	401	1	ASSY_CORGL	O85176	corynebacte
119	20	100.0	401	1	ASSY_STAAM	Q99vc7	staphylococ
120	20	100.0	401	1	ASSY_STAAW	Q8nxf2	staphylococ
121	20	100.0	401	1	ASSY_STAEP	Q8cpu3	staphylococ
122	20	100.0	401	1	ASSY_SYNEL	Q8dky7	synechococc
123	20	100.0	401	1	ASSY_XYLFA	Q9pem9	xylella fas
124	20	100.0	401	1	ASSY_XYLFT	P59606	xylella fas
125	20	100.0	402	1	ASSY_DEIRA	Q9rwj4	deinococcus
126	20	100.0	403	1	ASSY_BACSU	O34347	bacillus su
127	20	100.0	403	1	ASSY_BUCAI	P57158	buchnera ap
128	20	100.0	403	1	ASSY_LEPIN	Q8eyp7	leptospira
129	20	100.0	404	1	ASSY_LISIN	Q929s9	listeria in
130	20	100.0	404	1	ASSY_LISMO	Q8y5h2	listeria mo
131	20	100.0	404	1	ASSY_NITEU	Q82up5	nitrosomona

132	20	100.0	404	1	ASSY_STRAW	Q827z1	streptomyce
133	20	100.0	404	1	ASSY_VIBCH	Q9knt8	vibrio chol
134	20	100.0	404	1	ASSY_VIBPA	P59605	vibrio para
135	20	100.0	405	1	ASSY_CLOPE	Q8xmj7	clostridium
136	20	100.0	405	1	ASSY_PSEAE	Q9hy84	pseudomonas
137	20	100.0	405	1	ASSY_PSEPK	P59604	pseudomonas
138	20	100.0	405	1	ASSY_PSESM	Q87xm3	pseudomonas
139	20	100.0	405	1	ASSY_RHIME	Q92l73	rhizobium m
140	20	100.0	405	1	ASSY_THEAC	Q9hkf1	thermoplasm
141	20	100.0	405	1	ASSY_VIBVU	Q8dcn0	vibrio vuln
142	20	100.0	406	1	ASSY_BRUME	Q8yek8	brucella me
143	20	100.0	406	1	ASSY_BRUSU	Q8g376	brucella su
144	20	100.0	406	1	ASSY_CAMJE	Q9phk7	campylobact
145	20	100.0	406	1	ASSY_SULSO	Q9ux31	sulfolobus
146	20	100.0	407	1	ASSY_AGRT5	Q8uc31	agrobacteri
147	20	100.0	407	1	ASSY_RHILO	Q98e81	rhizobium l
148	20	100.0	407	1	ASSY_SHEON	Q8ek28	shewanella
149	20	100.0	408	1	ASSY_CAUCR	Q9abul	caulobacter
150	20	100.0	408	1	ASSY_HELMO	Q8gdu2	heliobacill

#### ALIGNMENTS

##### RESULT 1

##### ALLS\_MANSE

ID ALLS\_MANSE STANDARD; PRT; 15 AA.  
AC P42559;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Allatostatin (Mas-AS).  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;  
OC Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Head;  
RX MEDLINE=92052112; PubMed=1946359;  
RA Kramer S.J., Toschi A., Miller C.A., Kataoka H., Quistad G.B.,  
RA Li J.P., Carney R.L., Schooley D.A.;  
RT "Identification of an allatostatin from the tobacco hornworm Manduca  
RT sexta.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).  
CC -!- FUNCTION: Strongly inhibits juvenile hormone biosynthesis in vitro  
CC by the corpora allata from fifth-stadium larvae and adult females.  
CC -!- SIMILARITY: Belongs to the allatostatin family.  
DR PIR; A61612; A61612.  
KW Neuropeptide; Pyrrolidone carboxylic acid.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 15 AA; 1908 MW; 1605B77CDEBC838E CRC64;

Query Match 100.0%; Score 20; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



QY 1 QVRF 4  
 ||||  
 Db 1 QVRF 4

RESULT 2

YM26\_MARPO

ID YM26\_MARPO STANDARD; PRT; 63 AA.  
 AC P38468;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 7.1 kDa protein in NAD2 3'region (ORF 63).  
 GN YMF26.  
 OS Marchantia polymorpha (Liverwort).  
 OG Mitochondrion.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
 OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiineae;  
 OC Marchantiaceae; Marchantia.  
 OX NCBI\_TaxID=3197;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92114051; PubMed=1731062;  
 RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,  
 RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohyama K.;  
 RT "Gene organization deduced from the complete sequence of liverwort  
 RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant  
 RT mitochondrial genome.";  
 RL J. Mol. Biol. 223:1-7(1992).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M68929; AAC09400.1; -.  
 DR PIR; S25962; S25962.  
 KW Mitochondrion; Hypothetical protein.  
 SQ SEQUENCE 63 AA; 7143 MW; 607FCCB01A9FDA2A CRC64;

Query Match 100.0%; Score 20; DB 1; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 27;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QVRF 4  
 ||||  
 Db 6 QVRF 9

RESULT 3

Y65\_BPT3

ID Y65\_BPT3 STANDARD; PRT; 81 AA.  
 AC P20329;

```
Run on:      April  7, 2004, 19:09:21 ; Search time 16.1345 Seconds
              (without alignments)
              78.222 Million cell updates/sec
```

```
Title:      US-10-030-735-53
Perfect score: 20
Sequence:    1 QVRF 4
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 150 summaries
```

```
Database : SPTREMBL_25:*
1:  sp_archea:*
2:  sp_bacteria:*
3:  sp_fungi:*
4:  sp_human:*
5:  sp_invertebrate:*
6:  sp_mammal:*
7:  sp_mhc:*
8:  sp_organelle:*
9:  sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No..							

1	20	100.0	26	10	Q38749	Q38749 asparagus o
2	20	100.0	51	2	Q9S6X9	Q9s6x9 mycobacteri
3	20	100.0	54	16	Q7UR87	Q7ur87 rhodopirell
4	20	100.0	57	16	Q89LY6	Q89ly6 bradyrhizob
5	20	100.0	60	12	Q91B97	Q91b97 spodoptera
6	20	100.0	60	16	Q8PM94	Q8pm94 xanthomonas
7	20	100.0	62	16	Q8PAJ2	Q8paj2 xanthomonas
8	20	100.0	64	2	Q9JP69	Q9jp69 escherichia
9	20	100.0	66	7	Q9TPD1	Q9tpd1 callicebus
10	20	100.0	66	10	Q39945	Q39945 helianthus
11	20	100.0	70	10	Q84NC0	Q84nc0 oryza sativ
12	20	100.0	71	16	Q9JUN9	Q9jun9 neisseria m
13	20	100.0	73	16	Q8YTA8	Q8yta8 anabaena sp
14	20	100.0	76	7	Q31028	Q31028 saguinus oe
15	20	100.0	76	7	Q31025	Q31025 saguinus oe
16	20	100.0	76	7	Q31024	Q31024 saguinus oe
17	20	100.0	76	7	Q31029	Q31029 saguinus oe
18	20	100.0	76	7	Q31027	Q31027 saguinus oe
19	20	100.0	76	7	Q31026	Q31026 saguinus oe
20	20	100.0	79	7	O19366	O19366 felis silve
21	20	100.0	79	7	O19368	O19368 felis silve
22	20	100.0	79	7	O19375	O19375 felis silve
23	20	100.0	79	7	O19371	O19371 felis silve
24	20	100.0	79	7	O19377	O19377 felis silve
25	20	100.0	79	7	O19376	O19376 felis silve
26	20	100.0	79	7	O19373	O19373 felis silve
27	20	100.0	79	7	Q8MHQ4	Q8mhq4 felis silve
28	20	100.0	79	7	O19365	O19365 felis silve
29	20	100.0	79	7	O19369	O19369 felis silve
30	20	100.0	79	7	O19372	O19372 felis silve
31	20	100.0	79	7	Q8MHQ5	Q8mhq5 felis silve
32	20	100.0	79	7	O19388	O19388 felis silve
33	20	100.0	79	7	O19374	O19374 felis silve
34	20	100.0	79	7	O19370	O19370 felis silve
35	20	100.0	79	7	O19367	O19367 felis silve
36	20	100.0	80	16	Q89I80	Q89i80 bradyrhizob
37	20	100.0	81	9	Q9T113	Q9t113 bacterioph
38	20	100.0	81	16	Q8DIR4	Q8dir4 synechococc
39	20	100.0	82	2	Q8VMX3	Q8vmx3 listeria mo
40	20	100.0	82	7	Q30678	Q30678 macaca mula
41	20	100.0	82	7	Q95599	Q95599 phasianus c
42	20	100.0	82	7	Q30598	Q30598 macaca mula
43	20	100.0	83	7	Q9TNX4	Q9tnx4 equus cabal
44	20	100.0	83	7	Q860S5	Q860s5 sus scrofa
45	20	100.0	84	9	Q858L3	Q858l3 yersinia pe
46	20	100.0	85	2	Q8GJL9	Q8gjl9 synechococc
47	20	100.0	85	8	Q85G07	Q85g07 cyanidiosch
48	20	100.0	85	17	Q8TP81	Q8tp81 methanosarc
49	20	100.0	86	16	Q8DCD9	Q8dcd9 vibrio vuln
50	20	100.0	89	5	Q9XVX1	Q9xvx1 caenorhabdi
51	20	100.0	89	7	Q30859	Q30859 otolemur ga
52	20	100.0	89	7	Q30857	Q30857 otolemur ga
53	20	100.0	89	7	Q9TPD5	Q9tpd5 callicebus
54	20	100.0	89	7	Q30855	Q30855 otolemur ga
55	20	100.0	89	7	Q30854	Q30854 otolemur ga
56	20	100.0	89	7	Q30861	Q30861 otolemur ga
57	20	100.0	91	16	Q7URR6	Q7urr6 rhodopirell

58	20	100.0	92	16	Q8XEC0	Q8xec0 escherichia
59	20	100.0	94	2	Q842G6	Q842g6 pseudomonas
60	20	100.0	95	16	Q45063	Q45063 bacillus su
61	20	100.0	96	12	Q84679	Q84679 paramecium
62	20	100.0	97	2	Q7WYF4	Q7wyf4 pseudomonas
63	20	100.0	100	5	Q8IFB8	Q8ifb8 trypanosoma
64	20	100.0	101	16	Q9KSH5	Q9ksh5 vibrio chol
65	20	100.0	101	16	Q927F8	Q927f8 listeria in
66	20	100.0	101	16	Q8Y3Z5	Q8y3z5 listeria mo
67	20	100.0	101	16	Q87LH5	Q87lh5 vibrio para
68	20	100.0	102	4	Q15494	Q15494 homo sapien
69	20	100.0	102	10	Q8LCQ5	Q8lqc5 arabidopsis
70	20	100.0	102	16	Q9JUB1	Q9jub1 neisseria m
71	20	100.0	102	16	Q92AT8	Q92at8 listeria in
72	20	100.0	102	16	Q8Y6G7	Q8y6g7 listeria mo
73	20	100.0	102	16	Q82R15	Q82r15 streptomyce
74	20	100.0	103	16	Q8FHD8	Q8fhd8 escherichia
75	20	100.0	104	16	Q8D4B8	Q8d4b8 vibrio vuln
76	20	100.0	105	2	Q9EV95	Q9ev95 clostridium
77	20	100.0	105	16	Q836U5	Q836u5 enterococcu
78	20	100.0	106	5	Q7YWZ5	Q7y wz5 caenorhabdi
79	20	100.0	106	16	Q81YU2	Q81yu2 bacillus an
80	20	100.0	106	16	Q81I76	Q81i76 bacillus ce
81	20	100.0	107	3	Q05401	Q05401 saccharomyc
82	20	100.0	108	2	Q7WYE0	Q7wye0 acinetobact
83	20	100.0	108	16	Q836U0	Q836u0 enterococcu
84	20	100.0	109	16	Q836F5	Q836f5 enterococcu
85	20	100.0	110	10	Q7XVB9	Q7xvb9 oryza sativ
86	20	100.0	119	16	Q8DF56	Q8df56 vibrio vuln
87	20	100.0	119	16	Q87RX9	Q87rx9 vibrio para
88	20	100.0	120	16	Q8UJW8	Q8ujw8 agrobacteri
89	20	100.0	123	5	Q8IPA0	Q8ipa0 drosophila
90	20	100.0	124	3	Q06724	Q06724 saccharomyc
91	20	100.0	124	16	Q8F3H9	Q8f3h9 leptospira
92	20	100.0	125	5	Q26115	Q26115 pseudaletia
93	20	100.0	125	5	Q868F8	Q868f8 spodoptera
94	20	100.0	125	10	Q9LQT3	Q9lqt3 arabidopsis
95	20	100.0	128	16	Q9PK81	Q9pk81 chlamydia m
96	20	100.0	131	16	P73579	P73579 synechocyst
97	20	100.0	132	12	Q9QGH5	Q9qgh5 bean golden
98	20	100.0	132	12	O71203	O71203 potato yell
99	20	100.0	132	12	Q67581	Q67581 bean golden
100	20	100.0	132	12	Q67588	Q67588 bean golden
101	20	100.0	132	12	Q8JMJ6	Q8j mj6 potato yell
102	20	100.0	132	12	Q96702	Q96702 cabbage lea
103	20	100.0	132	12	Q80PA6	Q80pa6 bean golden
104	20	100.0	132	16	Q87KT2	Q87kt2 vibrio para
105	20	100.0	134	12	Q8V6Z6	Q8v6z6 african cas
106	20	100.0	134	12	Q8QUW2	Q8quw2 african cas
107	20	100.0	134	12	Q9JEA4	Q9jea4 cassava gem
108	20	100.0	134	12	Q9JEB1	Q9jeb1 cassava gem
109	20	100.0	134	12	Q8B657	Q8b657 african cas
110	20	100.0	135	6	Q9GL46	Q9gl46 sus scrofa
111	20	100.0	136	15	O11914	O11914 human immun
112	20	100.0	136	16	Q9K9K4	Q9k9k4 bacillus ha
113	20	100.0	138	6	O97671	O97671 oryctolagus
114	20	100.0	138	16	Q8RHG8	Q8rhg8 fusobacteri

115	20	100.0	139	4	Q8N939	Q8n939 homo sapien
116	20	100.0	140	2	Q84HI6	Q84hi6 azoarcus ev
117	20	100.0	140	5	P91191	P91191 caenorhabdi
118	20	100.0	140	11	Q921R2	Q921r2 mus musculu
119	20	100.0	141	2	Q7WZ61	Q7wz61 nonomuraea
120	20	100.0	144	2	P96031	P96031 shigella so
121	20	100.0	144	2	O82910	O82910 escherichia
122	20	100.0	144	2	Q9Z4D3	Q9z4d3 escherichia
123	20	100.0	144	2	Q849V0	Q849v0 escherichia
124	20	100.0	144	5	Q7YTD9	Q7ytd9 entamoeba m
125	20	100.0	144	16	Q83CQ5	Q83cq5 coxiella bu
126	20	100.0	144	16	Q7W0R3	Q7w0r3 bordetella
127	20	100.0	145	2	Q9ZGR7	Q9zgr7 escherichia
128	20	100.0	145	5	Q86E87	Q86e87 schistosoma
129	20	100.0	145	16	Q83HU9	Q83hu9 tropheryma
130	20	100.0	145	16	Q83GD2	Q83gd2 tropheryma
131	20	100.0	146	16	Q8ZJT5	Q8zjt5 yersinia pe
132	20	100.0	147	5	Q7YTE1	Q7yte1 entamoeba t
133	20	100.0	147	16	Q8UK50	Q8uk50 agrobacteri
134	20	100.0	148	5	Q8ITB0	Q8itb0 aequipecten
135	20	100.0	148	16	Q8YTV2	Q8ytv2 anabaena sp
136	20	100.0	150	10	Q84RW6	Q84rw6 oryza sativ
137	20	100.0	150	15	O56350	O56350 human immun
138	20	100.0	150	16	Q7WL25	Q7wl25 bordetella
139	20	100.0	150	16	Q7W7N7	Q7w7n7 bordetella
140	20	100.0	150	16	Q7VV99	Q7vv99 bordetella
141	20	100.0	151	5	Q8MUR2	Q8mur2 choristoneu
142	20	100.0	151	5	Q962R6	Q962r6 spodoptera
143	20	100.0	151	5	Q8ISP3	Q8isp3 branchiosto
144	20	100.0	151	5	Q8I7U0	Q8i7u0 plutella xy
145	20	100.0	151	5	Q8I7D6	Q8i7d6 ciona intes
146	20	100.0	151	11	Q9WVH0	Q9wvh0 cricetulus
147	20	100.0	151	13	Q9DFR6	Q9dfr6 gillichthys
148	20	100.0	151	13	Q90YR1	Q90yr1 ictalurus p
149	20	100.0	151	13	Q7SZT6	Q7szt6 xenopus lae
150	20	100.0	154	11	Q8BZ86	Q8bz86 mus musculu

# ALIGNMENTS

## RESULT 1

Q38749

ID Q38749 PRELIMINARY; PRT; 26 AA.

AC Q38749;

DT 01-NOV-1996 (TrEMBLrel. 01, Created)

DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Asp22 MADS box (Fragment).

OS Asparagus officinalis (Garden asparagus).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Asparagaceae;

OC Asparagus.

OX NCBI\_TaxID=4686;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=asp22;

RA Miller H.G., Kocher T.D., Loy B.;  
 RT "New MADS box domains in *Asparagus officinalis* L.";  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION  
 CC FACTORS.  
 DR EMBL; U07332; AAA18769.1; -.  
 DR GO; GO:0005634; C:nucleus; IEA.  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro; IPR000886; ER\_target\_S.  
 DR InterPro; IPR002100; TF\_MADSbox.  
 DR Pfam; PF00319; SRF-TF; 1.  
 DR SMART; SM00432; MADS; 1.  
 DR PROSITE; PS00014; ER\_TARGET; 1.  
 DR PROSITE; PS50066; MADS\_BOX\_2; 1.  
 KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.  
 FT NON\_TER 1 1  
 FT NON\_TER 26 26  
 SQ SEQUENCE 26 AA; 3120 MW; 020D523886B81B33 CRC64;

Query Match 100.0%; Score 20; DB 10; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVRF 4  
 ||||  
 Db 9 QVRF 12

## RESULT 2

Q9S6X9

ID Q9S6X9 PRELIMINARY; PRT; 51 AA.  
 AC Q9S6X9;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)  
 DE PGB14T-B precursor (Fragment).  
 OS *Mycobacterium tuberculosis*.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37Rv;  
 RX MEDLINE=98304086; PubMed=9639933;  
 RA Chubb A.J., Woodman Z.L., da Silva Tatley F.M.P.R., Hoffmann H.J.,  
 RA Scholle R.R., Ehlers M.R.W.;  
 RT "Identification of *Mycobacterium tuberculosis* signal sequences that  
 RT direct the export of a leaderless beta-lactamase gene product in  
 RT *Escherichia coli*.";  
 RL Microbiology 144:1619-1629(1998).  
 DR EMBL; AF017102; AAC26397.1; -.  
 DR InterPro; IPR003399; Mce.  
 DR Pfam; PF02470; mce; 1.  
 KW Signal.

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:16:36 ; Search time 6.28571 Seconds  
 (without alignments)  
 32.853 Million cell updates/sec

Title: US-10-030-735-54  
 Perfect score: 21  
 Sequence: 1 DVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	% Query Match	Length	DB	ID	Description
1	21	100.0	13	2	US-08-616-844-49	Sequence 49, Appl
2	21	100.0	13	2	US-08-599-654-49	Sequence 49, Appl
3	21	100.0	13	3	US-08-944-868A-49	Sequence 49, Appl
4	21	100.0	13	3	US-08-944-423A-49	Sequence 49, Appl
5	21	100.0	13	3	US-08-944-496-49	Sequence 49, Appl
6	21	100.0	13	4	US-08-908-371B-5	Sequence 5, Appli
7	21	100.0	17	3	US-08-996-679-63	Sequence 63, Appl
8	21	100.0	17	3	US-08-939-853A-14	Sequence 14, Appl
9	21	100.0	17	3	US-09-115-395-23	Sequence 23, Appl
10	21	100.0	17	3	US-09-113-977C-68	Sequence 68, Appl
11	21	100.0	17	4	US-09-250-059-54	Sequence 54, Appl

12	21	100.0	17	4	US-09-248-074-54	Sequence 54, Appl
13	21	100.0	17	4	US-09-187-859-50	Sequence 50, Appl
14	21	100.0	17	4	US-09-458-870-54	Sequence 54, Appl
15	21	100.0	17	4	US-09-351-048A-68	Sequence 68, Appl
16	21	100.0	17	4	US-09-839-542B-50	Sequence 50, Appl
17	21	100.0	17	4	US-09-535-852-1935	Sequence 1935, Ap
18	21	100.0	20	4	US-09-220-415-34	Sequence 34, Appl
19	21	100.0	20	4	US-09-675-776-34	Sequence 34, Appl
20	21	100.0	20	4	US-09-676-739-34	Sequence 34, Appl
21	21	100.0	28	3	US-08-921-887-31	Sequence 31, Appl
22	21	100.0	28	4	US-09-149-476-662	Sequence 662, App
23	21	100.0	31	1	US-08-190-802A-229	Sequence 229, App
24	21	100.0	31	1	US-08-324-301-12	Sequence 12, Appl
25	21	100.0	31	3	US-08-477-346-229	Sequence 229, App
26	21	100.0	31	4	US-08-473-089-229	Sequence 229, App
27	21	100.0	31	4	US-08-487-072A-229	Sequence 229, App
28	21	100.0	33	4	US-08-525-539A-18	Sequence 18, Appl
29	21	100.0	48	4	US-09-079-030-102	Sequence 102, App
30	21	100.0	63	4	US-09-134-001C-2876	Sequence 2876, Ap
31	21	100.0	63	4	US-09-663-600A-88	Sequence 88, Appl
32	21	100.0	74	4	US-09-198-452A-1267	Sequence 1267, Ap
33	21	100.0	80	4	US-09-630-273-2	Sequence 2, Appli
34	21	100.0	81	4	US-09-333-809-193	Sequence 193, App
35	21	100.0	83	4	US-09-333-809-94	Sequence 94, Appl
36	21	100.0	86	4	US-09-448-806C-3	Sequence 3, Appli
37	21	100.0	88	2	US-08-327-362-2	Sequence 2, Appli
38	21	100.0	88	2	US-08-709-924-24	Sequence 24, Appl
39	21	100.0	88	2	US-08-709-925-24	Sequence 24, Appl
40	21	100.0	88	3	US-09-158-565-2	Sequence 2, Appli
41	21	100.0	88	4	US-08-709-948-24	Sequence 24, Appl
42	21	100.0	88	4	US-09-220-415-24	Sequence 24, Appl
43	21	100.0	88	4	US-09-675-776-24	Sequence 24, Appl
44	21	100.0	88	4	US-09-676-739-24	Sequence 24, Appl
45	21	100.0	91	4	US-09-252-991A-31622	Sequence 31622, A
46	21	100.0	93	3	US-08-918-288-70	Sequence 70, Appl
47	21	100.0	93	3	US-09-282-357-70	Sequence 70, Appl
48	21	100.0	93	3	US-08-975-040-24	Sequence 24, Appl
49	21	100.0	93	4	US-09-134-000C-5269	Sequence 5269, Ap
50	21	100.0	93	4	US-09-050-739-141	Sequence 141, App
51	21	100.0	98	2	US-08-709-924-23	Sequence 23, Appl
52	21	100.0	98	2	US-08-709-925-23	Sequence 23, Appl
53	21	100.0	98	4	US-08-709-948-23	Sequence 23, Appl
54	21	100.0	98	4	US-09-220-415-23	Sequence 23, Appl
55	21	100.0	98	4	US-09-675-776-23	Sequence 23, Appl
56	21	100.0	98	4	US-09-676-739-23	Sequence 23, Appl
57	21	100.0	102	3	US-08-946-329A-102	Sequence 102, App
58	21	100.0	109	4	US-09-540-236-3197	Sequence 3197, Ap
59	21	100.0	114	1	US-08-425-673-7	Sequence 7, Appli
60	21	100.0	114	1	US-08-425-673-8	Sequence 8, Appli
61	21	100.0	114	1	US-08-425-673-9	Sequence 9, Appli
62	21	100.0	114	3	US-08-918-288-69	Sequence 69, Appl
63	21	100.0	114	3	US-08-918-288-71	Sequence 71, Appl
64	21	100.0	114	3	US-09-282-357-69	Sequence 69, Appl
65	21	100.0	114	3	US-09-282-357-71	Sequence 71, Appl
66	21	100.0	117	1	US-08-425-673-11	Sequence 11, Appl
67	21	100.0	117	1	US-08-425-673-12	Sequence 12, Appl
68	21	100.0	122	4	US-09-630-273-1	Sequence 1, Appli



69	21	100.0	130	4	US-09-134-000C-5192	Sequence 5192, Ap
70	21	100.0	145	1	US-08-425-673-1	Sequence 1, Appli
71	21	100.0	145	1	US-08-425-673-2	Sequence 2, Appli
72	21	100.0	145	1	US-08-425-673-10	Sequence 10, Appl
73	21	100.0	145	1	US-08-298-189B-1	Sequence 1, Appli
74	21	100.0	145	1	US-08-475-213-10	Sequence 10, Appl
75	21	100.0	145	2	US-08-395-238-2	Sequence 2, Appli
76	21	100.0	145	3	US-09-142-320-4	Sequence 4, Appli
77	21	100.0	145	3	US-09-142-320-11	Sequence 11, Appl
78	21	100.0	145	3	US-09-142-320-12	Sequence 12, Appl
79	21	100.0	145	3	US-09-142-320-13	Sequence 13, Appl
80	21	100.0	145	3	US-09-142-320-14	Sequence 14, Appl
81	21	100.0	145	3	US-09-142-320-15	Sequence 15, Appl
82	21	100.0	145	3	US-09-142-320-16	Sequence 16, Appl
83	21	100.0	145	3	US-08-918-288-68	Sequence 68, Appl
84	21	100.0	145	3	US-09-282-357-68	Sequence 68, Appl
85	21	100.0	145	4	US-08-908-371B-1	Sequence 1, Appli
86	21	100.0	145	4	US-09-760-294-4	Sequence 4, Appli
87	21	100.0	145	4	US-09-760-294-11	Sequence 11, Appl
88	21	100.0	145	4	US-09-760-294-12	Sequence 12, Appl
89	21	100.0	145	4	US-09-760-294-13	Sequence 13, Appl
90	21	100.0	145	4	US-09-760-294-14	Sequence 14, Appl
91	21	100.0	145	4	US-09-760-294-15	Sequence 15, Appl
92	21	100.0	145	4	US-09-760-294-16	Sequence 16, Appl
93	21	100.0	151	4	US-09-059-625-59	Sequence 59, Appl
94	21	100.0	151	4	US-09-480-297A-6	Sequence 6, Appli
95	21	100.0	155	4	US-09-853-832-14	Sequence 14, Appl
96	21	100.0	158	4	US-09-059-625-1	Sequence 1, Appli
97	21	100.0	160	4	US-09-231-788-12	Sequence 12, Appl
98	21	100.0	165	2	US-08-709-924-2	Sequence 2, Appli
99	21	100.0	165	2	US-08-709-925-2	Sequence 2, Appli
100	21	100.0	165	4	US-08-709-948-2	Sequence 2, Appli
101	21	100.0	165	4	US-09-220-415-2	Sequence 2, Appli
102	21	100.0	165	4	US-09-675-776-2	Sequence 2, Appli
103	21	100.0	165	4	US-09-676-739-2	Sequence 2, Appli
104	21	100.0	166	4	US-09-107-532A-5004	Sequence 5004, Ap
105	21	100.0	174	4	US-09-149-476-526	Sequence 526, App
106	21	100.0	181	3	US-08-918-288-36	Sequence 36, Appl
107	21	100.0	181	3	US-09-282-357-36	Sequence 36, Appl
108	21	100.0	181	4	US-09-059-625-66	Sequence 66, Appl
109	21	100.0	183	1	US-08-602-010A-4	Sequence 4, Appli
110	21	100.0	183	1	US-08-680-726A-4	Sequence 4, Appli
111	21	100.0	183	3	US-09-092-409-4	Sequence 4, Appli
112	21	100.0	188	4	US-09-489-039A-7222	Sequence 7222, Ap
113	21	100.0	198	1	US-08-052-681-1	Sequence 1, Appli
114	21	100.0	201	4	US-09-543-681A-7854	Sequence 7854, Ap
115	21	100.0	202	4	US-09-480-297A-8	Sequence 8, Appli
116	21	100.0	202	4	US-09-747-259-8	Sequence 8, Appli
117	21	100.0	202	4	US-09-816-744-8	Sequence 8, Appli
118	21	100.0	203	4	US-09-059-625-85	Sequence 85, Appl
119	21	100.0	203	4	US-08-311-731A-31	Sequence 31, Appl
120	21	100.0	204	4	US-09-059-625-35	Sequence 35, Appl
121	21	100.0	204	4	US-09-059-625-49	Sequence 49, Appl
122	21	100.0	204	4	US-09-059-625-54	Sequence 54, Appl
123	21	100.0	206	4	US-09-059-625-90	Sequence 90, Appl
124	21	100.0	208	4	US-09-059-625-63	Sequence 63, Appl
125	21	100.0	209	4	US-09-059-625-86	Sequence 86, Appl

126	21	100.0	209	4	US-09-059-625-87	Sequence 87, Appl
127	21	100.0	212	4	US-09-059-625-58	Sequence 58, Appl
128	21	100.0	212	4	US-09-059-625-65	Sequence 65, Appl
129	21	100.0	212	4	US-09-059-625-69	Sequence 69, Appl
130	21	100.0	212	4	US-09-059-625-70	Sequence 70, Appl
131	21	100.0	213	4	US-09-328-352-5631	Sequence 5631, Ap
132	21	100.0	220	4	US-09-252-991A-25919	Sequence 25919, A
133	21	100.0	222	4	US-09-328-352-7790	Sequence 7790, Ap
134	21	100.0	234	3	US-08-918-288-6	Sequence 6, Appli
135	21	100.0	234	3	US-08-918-288-9	Sequence 9, Appli
136	21	100.0	234	3	US-08-918-288-21	Sequence 21, Appl
137	21	100.0	234	3	US-08-918-288-24	Sequence 24, Appl
138	21	100.0	234	3	US-09-282-357-6	Sequence 6, Appli
139	21	100.0	234	3	US-09-282-357-9	Sequence 9, Appli
140	21	100.0	234	3	US-09-282-357-21	Sequence 21, Appl
141	21	100.0	234	3	US-09-282-357-24	Sequence 24, Appl
142	21	100.0	237	3	US-08-918-288-15	Sequence 15, Appl
143	21	100.0	237	3	US-08-918-288-18	Sequence 18, Appl
144	21	100.0	237	3	US-09-282-357-15	Sequence 15, Appl
145	21	100.0	237	3	US-09-282-357-18	Sequence 18, Appl
146	21	100.0	237	4	US-09-543-681A-5265	Sequence 5265, Ap
147	21	100.0	238	3	US-08-634-475-6	Sequence 6, Appli
148	21	100.0	238	4	US-09-709-791-6	Sequence 6, Appli
149	21	100.0	238	4	US-09-874-585D-44	Sequence 44, Appl
150	21	100.0	245	4	US-09-059-625-73	Sequence 73, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-616-844-49

; Sequence 49, Application US/08616844

; Patent No. 5849578

; GENERAL INFORMATION:

; APPLICANT: FALB, DEAN A.

; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE

; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE

; NUMBER OF SEQUENCES: 54

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: PENNIE & EDMONDS

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/616,844

; FILING DATE: 15-MAR-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/599,654

; FILING DATE: 09-FEB-1996  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/485,573  
 ; FILING DATE: 07-JUN-1995  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/386,844  
 ; FILING DATE: 10-FEB-1995  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: CORUZZI, LAURA A.  
 ; REGISTRATION NUMBER: 30,742  
 ; REFERENCE/DOCKET NUMBER: 7853-053  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (212) 790-9090  
 ; TELEFAX: (212) 869-8864  
 ; TELEX: 66141 PENNIE  
 ; INFORMATION FOR SEQ ID NO: 49:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 13 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: unknown  
 ; MOLECULE TYPE: peptide  
 US-08-616-844-49

Query Match 100.0%; Score 21; DB 2; Length 13;  
 Best Local Similarity 100.0%; Pred. No. 36;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVRF 4  
 ||||  
 Db 9 DVRF 12

# RESULT 2

US-08-599-654-49  
 ; Sequence 49, Application US/08599654  
 ; Patent No. 5882925  
 ; GENERAL INFORMATION:  
 ; APPLICANT: FALB, DEAN A  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE  
 ; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE  
 ; NUMBER OF SEQUENCES: 54  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: PENNIE & EDMONDS  
 ; STREET: 1155 Avenue of the Americas  
 ; CITY: New York  
 ; STATE: New York  
 ; COUNTRY: USA  
 ; ZIP: 10036-2711  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/599,654  
 ; FILING DATE: 09-FEB-1996

OM protein - protein search, using sw model

Run on: April 7, 2004, 22:05:26 ; Search time 16.8403 Seconds  
 (without alignments)  
 62.382 Million cell updates/sec

Title: US-10-030-735-54  
 Perfect score: 21  
 Sequence: 1 DVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Match	Length	ID	Description
No.					

1	21	100.0	10	15	US-10-149-138-4326	Sequence 4326, Ap
2	21	100.0	12	14	US-10-185-050-184	Sequence 184, App
3	21	100.0	13	9	US-09-371-900-49	Sequence 49, Appl
4	21	100.0	13	9	US-09-924-417-15	Sequence 15, Appl
5	21	100.0	13	14	US-10-186-950-49	Sequence 49, Appl
6	21	100.0	17	9	US-09-769-145-54	Sequence 54, Appl
7	21	100.0	17	10	US-09-778-026-14	Sequence 14, Appl
8	21	100.0	17	12	US-10-412-701-54	Sequence 54, Appl
9	21	100.0	17	14	US-10-006-869-50	Sequence 50, Appl
10	21	100.0	17	14	US-10-193-653-68	Sequence 68, Appl
11	21	100.0	17	15	US-10-395-032-50	Sequence 50, Appl
12	21	100.0	17	15	US-10-425-557-54	Sequence 54, Appl
13	21	100.0	28	10	US-09-809-391-662	Sequence 662, App
14	21	100.0	28	10	US-09-882-171-662	Sequence 662, App
15	21	100.0	33	9	US-09-956-206A-18	Sequence 18, Appl
16	21	100.0	35	14	US-10-028-075B-169	Sequence 169, App
17	21	100.0	35	14	US-10-029-206A-169	Sequence 169, App
18	21	100.0	38	14	US-10-028-075B-35	Sequence 35, Appl
19	21	100.0	38	14	US-10-029-206A-35	Sequence 35, Appl
20	21	100.0	38	15	US-10-409-630-24	Sequence 24, Appl
21	21	100.0	38	15	US-10-409-027-24	Sequence 24, Appl
22	21	100.0	38	15	US-10-409-032-24	Sequence 24, Appl
23	21	100.0	38	15	US-10-409-642-24	Sequence 24, Appl
24	21	100.0	38	15	US-10-409-654-24	Sequence 24, Appl
25	21	100.0	38	15	US-10-409-659-24	Sequence 24, Appl
26	21	100.0	38	15	US-10-409-671-24	Sequence 24, Appl
27	21	100.0	38	15	US-10-409-657-24	Sequence 24, Appl
28	21	100.0	42	12	US-10-424-599-257922	Sequence 257922,
29	21	100.0	47	12	US-10-424-599-209542	Sequence 209542,
30	21	100.0	49	12	US-10-424-599-152393	Sequence 152393,
31	21	100.0	49	12	US-10-424-599-200060	Sequence 200060,
32	21	100.0	51	14	US-10-106-698-7534	Sequence 7534, Ap
33	21	100.0	53	12	US-10-424-599-255848	Sequence 255848,
34	21	100.0	54	12	US-10-424-599-156597	Sequence 156597,
35	21	100.0	57	12	US-10-424-599-274982	Sequence 274982,
36	21	100.0	58	12	US-10-424-599-181046	Sequence 181046,
37	21	100.0	59	11	US-09-864-408A-4292	Sequence 4292, Ap
38	21	100.0	61	12	US-10-424-599-153530	Sequence 153530,
39	21	100.0	61	14	US-10-193-477-11	Sequence 11, Appl
40	21	100.0	63	12	US-10-424-599-175392	Sequence 175392,
41	21	100.0	63	14	US-10-319-763-88	Sequence 88, Appl
42	21	100.0	65	14	US-10-029-386-31245	Sequence 31245, A
43	21	100.0	66	12	US-10-424-599-237964	Sequence 237964,
44	21	100.0	66	12	US-10-424-599-251379	Sequence 251379,
45	21	100.0	66	14	US-10-029-386-28776	Sequence 28776, A
46	21	100.0	67	12	US-10-424-599-254168	Sequence 254168,
47	21	100.0	69	12	US-10-282-122A-62091	Sequence 62091, A
48	21	100.0	69	12	US-10-424-599-160561	Sequence 160561,
49	21	100.0	70	12	US-10-424-599-145504	Sequence 145504,
50	21	100.0	72	12	US-10-424-599-257899	Sequence 257899,
51	21	100.0	74	12	US-10-424-599-254011	Sequence 254011,
52	21	100.0	74	15	US-10-289-762-1267	Sequence 1267, Ap
53	21	100.0	77	9	US-09-864-761-47518	Sequence 47518, A
54	21	100.0	80	12	US-10-424-599-200903	Sequence 200903,
55	21	100.0	80	12	US-10-424-599-205012	Sequence 205012,
56	21	100.0	80	12	US-10-361-652-2	Sequence 2, Appli

57	21	100.0	80	14	US-10-159-354-2	Sequence 2, Appli
58	21	100.0	81	12	US-10-169-395-129	Sequence 129, App
59	21	100.0	82	15	US-10-463-190-127	Sequence 127, App
60	21	100.0	86	12	US-10-424-599-205029	Sequence 205029,
61	21	100.0	86	12	US-10-425-114-50429	Sequence 50429, A
62	21	100.0	87	9	US-09-864-761-41358	Sequence 41358, A
63	21	100.0	87	9	US-09-864-761-45638	Sequence 45638, A
64	21	100.0	87	9	US-09-864-761-47449	Sequence 47449, A
65	21	100.0	87	12	US-10-424-599-181276	Sequence 181276,
66	21	100.0	88	12	US-10-424-599-209115	Sequence 209115,
67	21	100.0	88	14	US-10-050-875-24	Sequence 24, Appl
68	21	100.0	91	14	US-10-156-761-11467	Sequence 11467, A
69	21	100.0	92	9	US-09-738-626-6354	Sequence 6354, Ap
70	21	100.0	93	9	US-09-791-171-141	Sequence 141, App
71	21	100.0	93	10	US-09-804-980-141	Sequence 141, App
72	21	100.0	95	14	US-10-156-761-12697	Sequence 12697, A
73	21	100.0	96	12	US-10-335-977-6940	Sequence 6940, Ap
74	21	100.0	98	14	US-10-050-875-23	Sequence 23, Appl
75	21	100.0	102	15	US-10-460-594-102	Sequence 102, App
76	21	100.0	106	12	US-10-424-599-211757	Sequence 211757,
77	21	100.0	107	11	US-09-927-876-33	Sequence 33, Appl
78	21	100.0	107	15	US-10-360-149-33	Sequence 33, Appl
79	21	100.0	109	11	US-09-927-876-38	Sequence 38, Appl
80	21	100.0	109	11	US-09-927-876-40	Sequence 40, Appl
81	21	100.0	109	15	US-10-360-149-38	Sequence 38, Appl
82	21	100.0	109	15	US-10-360-149-40	Sequence 40, Appl
83	21	100.0	111	12	US-10-335-977-6942	Sequence 6942, Ap
84	21	100.0	113	12	US-10-424-599-250872	Sequence 250872,
85	21	100.0	114	12	US-10-424-599-260121	Sequence 260121,
86	21	100.0	116	11	US-09-864-408A-502	Sequence 502, App
87	21	100.0	117	12	US-10-425-114-51344	Sequence 51344, A
88	21	100.0	121	11	US-09-864-408A-8078	Sequence 8078, Ap
89	21	100.0	121	12	US-10-282-122A-48956	Sequence 48956, A
90	21	100.0	121	15	US-10-360-101-251	Sequence 251, App
91	21	100.0	122	9	US-09-813-398-4	Sequence 4, Appli
92	21	100.0	122	12	US-10-361-652-1	Sequence 1, Appli
93	21	100.0	122	14	US-10-159-354-1	Sequence 1, Appli
94	21	100.0	123	12	US-10-424-599-272788	Sequence 272788,
95	21	100.0	126	14	US-10-106-698-4324	Sequence 4324, Ap
96	21	100.0	129	12	US-10-424-599-173652	Sequence 173652,
97	21	100.0	131	12	US-10-424-599-166673	Sequence 166673,
98	21	100.0	131	12	US-10-425-114-52463	Sequence 52463, A
99	21	100.0	137	12	US-10-424-599-187911	Sequence 187911,
100	21	100.0	138	14	US-10-128-714-3003	Sequence 3003, Ap
101	21	100.0	139	14	US-10-156-761-8842	Sequence 8842, Ap
102	21	100.0	141	9	US-09-813-398-3	Sequence 3, Appli
103	21	100.0	141	11	US-09-927-876-6	Sequence 6, Appli
104	21	100.0	141	15	US-10-360-149-6	Sequence 6, Appli
105	21	100.0	142	12	US-10-424-599-175120	Sequence 175120,
106	21	100.0	143	12	US-10-424-599-176306	Sequence 176306,
107	21	100.0	143	12	US-10-425-114-50524	Sequence 50524, A
108	21	100.0	145	9	US-09-760-294-4	Sequence 4, Appli
109	21	100.0	145	9	US-09-760-294-11	Sequence 11, Appl
110	21	100.0	145	9	US-09-760-294-12	Sequence 12, Appl
111	21	100.0	145	9	US-09-760-294-13	Sequence 13, Appl
112	21	100.0	145	9	US-09-760-294-14	Sequence 14, Appl
113	21	100.0	145	9	US-09-760-294-15	Sequence 15, Appl

114	21	100.0	145	9	US-09-760-294-16	Sequence 16, Appl
115	21	100.0	145	14	US-10-131-241-65	Sequence 65, Appl
116	21	100.0	145	15	US-10-360-101-214	Sequence 214, App
117	21	100.0	149	9	US-09-815-242-11049	Sequence 11049, A
118	21	100.0	149	12	US-10-282-122A-58213	Sequence 58213, A
119	21	100.0	150	12	US-10-282-122A-61330	Sequence 61330, A
120	21	100.0	151	14	US-10-366-791-6	Sequence 6, Appli
121	21	100.0	152	10	US-09-729-838-1	Sequence 1, Appli
122	21	100.0	153	10	US-09-764-891-4027	Sequence 4027, Ap
123	21	100.0	154	12	US-10-424-599-185164	Sequence 185164,
124	21	100.0	155	9	US-09-961-452-14	Sequence 14, Appl
125	21	100.0	155	9	US-09-961-453-14	Sequence 14, Appl
126	21	100.0	155	10	US-09-961-507-14	Sequence 14, Appl
127	21	100.0	155	10	US-09-961-637-14	Sequence 14, Appl
128	21	100.0	155	10	US-09-961-458-14	Sequence 14, Appl
129	21	100.0	155	10	US-09-453-976-14	Sequence 14, Appl
130	21	100.0	155	12	US-10-387-345-14	Sequence 14, Appl
131	21	100.0	155	12	US-10-335-977-5045	Sequence 5045, Ap
132	21	100.0	155	14	US-10-284-084-14	Sequence 14, Appl
133	21	100.0	155	14	US-10-409-783-14	Sequence 14, Appl
134	21	100.0	156	14	US-10-156-761-14283	Sequence 14283, A
135	21	100.0	160	9	US-09-731-816-4	Sequence 4, Appli
136	21	100.0	160	10	US-09-320-713-4	Sequence 4, Appli
137	21	100.0	160	14	US-10-153-770-4	Sequence 4, Appli
138	21	100.0	160	14	US-10-209-041-3	Sequence 3, Appli
139	21	100.0	160	14	US-10-277-726A-12	Sequence 12, Appl
140	21	100.0	160	14	US-10-397-282-4	Sequence 4, Appli
141	21	100.0	165	9	US-09-466-320-14	Sequence 14, Appl
142	21	100.0	165	9	US-09-915-676-1	Sequence 1, Appli
143	21	100.0	165	11	US-09-927-876-8	Sequence 8, Appli
144	21	100.0	165	12	US-10-424-599-276699	Sequence 276699,
145	21	100.0	165	12	US-10-411-037-72	Sequence 72, Appl
146	21	100.0	165	14	US-10-050-875-2	Sequence 2, Appli
147	21	100.0	165	14	US-10-128-714-8003	Sequence 8003, Ap
148	21	100.0	165	15	US-10-360-149-8	Sequence 8, Appli
149	21	100.0	166	11	US-09-927-876-13	Sequence 13, Appl
150	21	100.0	166	12	US-10-424-599-259479	Sequence 259479,

#### ALIGNMENTS

##### RESULT 1

US-10-149-138-4326

; Sequence 4326, Application US/10149138

; Publication No. US20040018971A1

##### ; GENERAL INFORMATION:

; APPLICANT: Fikes, John

; APPLICANT: Sette, Alessandro

; APPLICANT: Sidney, John

; APPLICANT: Southwood, Scott

; APPLICANT: Chesnut, Robert

; APPLICANT: Celis, Esteban

; APPLICANT: Keogh, Elissa

; TITLE OF INVENTION: Inducing Cellular Immune Responses to

; TITLE OF INVENTION: HER2/neu Using Peptide and Nucleic Acid Compositions

; FILE REFERENCE: 2060.0140001

; CURRENT APPLICATION NUMBER: US/10/149,138  
; CURRENT FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: PCT/US00/33591  
; PRIOR FILING DATE: 2000-12-11  
; PRIOR APPLICATION NUMBER: US 09/458,299  
; PRIOR FILING DATE: 1999-12-11  
; NUMBER OF SEQ ID NOS: 4641  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4326  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Artificial Peptide  
US-10-149-138-4326

Query Match 100.0%; Score 21; DB 15; Length 10;  
Best Local Similarity 100.0%; Pred. No. 46;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVRF 4  
|||  
Db 7 DVRF 10

RESULT 2

US-10-185-050-184

; Sequence 184, Application US/10185050  
; Publication No. US20030077577A1  
; GENERAL INFORMATION:  
; APPLICANT: Pirozzi, Gregorio  
; Kay, Brian K.  
; Fowlkes, Dana M.  
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL  
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING  
SAME

; NUMBER OF SEQUENCES: 233  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/185,050  
; FILING DATE: 28-Jun-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/826,516  
; FILING DATE: 03-Apr-1997  
; ATTORNEY/AGENT INFORMATION:



OM protein - protein search, using sw model

Run on: April 7, 2004, 19:11:45 ; Search time 5.10924 Seconds  
 (without alignments)  
 75.308 Million cell updates/sec

Title: US-10-030-735-54  
 Perfect score: 21  
 Sequence: 1 DVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	% Query		Match	Length	DB	ID	Description
	Score						
1	21	100.0	17	2	S59512	probable integrin	
2	21	100.0	56	2	H82806	hypothetical prote	
3	21	100.0	57	2	D35054	MHC class II histo	
4	21	100.0	82	2	I51137	MHC class II beta	
5	21	100.0	82	2	I50968	MHC class II beta	
6	21	100.0	82	2	I50984	MHC class II beta	
7	21	100.0	82	2	I50986	MHC class II beta	
8	21	100.0	82	2	I51148	MHC class II beta	
9	21	100.0	87	2	E86949	hypothetical prote	
10	21	100.0	92	2	T36134	hypothetical prote	
11	21	100.0	93	2	C70771	hypothetical prote	
12	21	100.0	95	2	H81119	conserved hypothet	
13	21	100.0	102	1	R3HS15	ribosomal protein	

14	21	100.0	103	2	B64334	hypothetical prote
15	21	100.0	107	2	T48703	hypothetical prote
16	21	100.0	109	2	G87531	hypothetical prote
17	21	100.0	109	2	A95061	BlpT protein, fusi
18	21	100.0	109	2	E97929	hypothetical prote
19	21	100.0	113	2	T41404	hypothetical prote
20	21	100.0	115	2	F90277	hypothetical prote
21	21	100.0	121	2	C42573	NADH2 dehydrogenas
22	21	100.0	121	2	T11226	NADH2 dehydrogenas
23	21	100.0	122	2	S78167	NADH2 dehydrogenas
24	21	100.0	123	2	S41361	genome polyprotein
25	21	100.0	124	2	C97842	hypothetical prote
26	21	100.0	125	2	G87491	NADH dehydrogenase
27	21	100.0	126	2	H90326	conserved hypothet
28	21	100.0	127	1	S19498	thioredoxin homolo
29	21	100.0	132	2	F87105	hypothetical prote
30	21	100.0	132	2	H75335	hypothetical prote
31	21	100.0	136	2	PQ0517	pol protein - ovin
32	21	100.0	136	2	AD1869	hypothetical prote
33	21	100.0	138	2	T17770	hypothetical prote
34	21	100.0	141	1	UTHUB	lutropin beta chai
35	21	100.0	145	2	S20227	cytochrome P450 2C
36	21	100.0	145	2	I37231	beta-gonadotropin
37	21	100.0	146	1	C64085	hypothetical prote
38	21	100.0	148	2	T48981	hypothetical prote
39	21	100.0	149	2	D64076	ribosomal protein
40	21	100.0	150	2	S26068	hypothetical prote
41	21	100.0	152	2	C71973	hypothetical prote
42	21	100.0	152	2	T36819	hypothetical prote
43	21	100.0	153	2	I39827	hypothetical prote
44	21	100.0	154	2	A45533	T-cell-stimulating
45	21	100.0	155	2	A64533	conserved hypothet
46	21	100.0	155	2	AF1598	conserved hypothet
47	21	100.0	155	2	AH1235	conserved hypothet
48	21	100.0	157	2	B70377	iron-sulfur cofact
49	21	100.0	158	2	B86658	hypothetical prote
50	21	100.0	158	2	S44948	lmbB1 protein - St
51	21	100.0	161	2	D69491	monoamine oxidase
52	21	100.0	161	2	AE1322	hypothetical prote
53	21	100.0	161	2	AF1693	hypothetical prote
54	21	100.0	165	1	KTHUB	choriogonadotropin
55	21	100.0	167	2	B89897	hypothetical prote
56	21	100.0	169	2	AD1481	different protein
57	21	100.0	171	2	E87512	dihydrofolate redu
58	21	100.0	171	2	AI0432	probable stringent
59	21	100.0	171	2	H87440	hypothetical prote
60	21	100.0	172	2	F65038	hypothetical prote
61	21	100.0	172	2	B91062	hypothetical prote
62	21	100.0	172	2	F85906	hypothetical prote
63	21	100.0	176	2	E87523	hypoxanthine-guani
64	21	100.0	176	2	E71371	hypothetical prote
65	21	100.0	178	2	G87543	hypothetical prote
66	21	100.0	179	2	B82143	conserved hypothet
67	21	100.0	180	2	T39395	hypothetical oligo
68	21	100.0	180	2	AC0197	conserved hypothet
69	21	100.0	187	2	S74014	hypothetical prote
70	21	100.0	191	2	H71370	hypothetical prote

71	21	100.0	193	2	C82683	disulfide oxidored
72	21	100.0	195	2	AH2774	hypothetical prote
73	21	100.0	196	2	A29279	conserved hypothet
74	21	100.0	196	2	A88209	protein K02A2.4 [i
75	21	100.0	199	2	T19584	hypothetical prote
76	21	100.0	200	2	A82741	riboflavin synthas
77	21	100.0	203	2	E87200	DNA repair protein
78	21	100.0	204	2	S39735	ywfB protein - Bac
79	21	100.0	206	2	S72567	hypothetical prote
80	21	100.0	207	2	F83815	hypothetical prote
81	21	100.0	207	2	C56658	24K protein GP24C
82	21	100.0	209	2	T17825	hypothetical prote
83	21	100.0	210	2	E75553	uridine kinase - D
84	21	100.0	210	2	AB3420	hypothetical phage
85	21	100.0	210	2	A84348	hypothetical prote
86	21	100.0	211	2	B84961	lipoate-protein li
87	21	100.0	211	2	B85249	hypothetical prote
88	21	100.0	212	2	B81088	2-dehydro-3-deoxy-
89	21	100.0	212	2	A96510	protein F27F5.25 [
90	21	100.0	215	2	F83442	probable transcrip
91	21	100.0	215	2	S75160	hypothetical prote
92	21	100.0	216	2	C83879	transcription regu
93	21	100.0	217	2	T48949	hypothetical prote
94	21	100.0	221	2	S50476	hypothetical prote
95	21	100.0	222	2	H87496	hypothetical prote
96	21	100.0	225	2	E95187	Bcl-2 family prote
97	21	100.0	225	2	F98054	conserved hypothet
98	21	100.0	225	2	T05463	hypothetical prote
99	21	100.0	227	2	E64830	cytidylate kinase
100	21	100.0	227	2	A99753	cytidylate kinase
101	21	100.0	227	2	G85616	cytidylate kinase
102	21	100.0	227	2	AB0614	cytidylate kinase
103	21	100.0	230	2	AI0169	cytidylate kinase
104	21	100.0	230	2	H95265	probable GntR-fami
105	21	100.0	230	2	A96565	hypothetical prote
106	21	100.0	235	2	AI0154	pseudouridylate sy
107	21	100.0	236	2	AE1917	ribose 5-phosphate
108	21	100.0	238	2	S34072	riboflavin synthas
109	21	100.0	238	2	S70195	kfiA protein - Esc
110	21	100.0	241	2	A64364	hypothetical prote
111	21	100.0	241	2	S34500	hypothetical prote
112	21	100.0	244	2	H89942	conserved hypothet
113	21	100.0	244	2	B85061	hypothetical prote
114	21	100.0	244	2	S30860	hypothetical prote
115	21	100.0	246	2	F69259	hypothetical prote
116	21	100.0	248	2	G75214	methyltransferase
117	21	100.0	249	2	E90102	TATA binding prote
118	21	100.0	249	2	E90137	TATA binding prote
119	21	100.0	249	2	G90118	TATA binding prote
120	21	100.0	251	2	S69021	hypothetical prote
121	21	100.0	251	2	H70738	hypothetical prote
122	21	100.0	252	1	E69839	conserved hypothet
123	21	100.0	252	2	T28194	hypothetical prote
124	21	100.0	254	2	A75433	hypothetical prote
125	21	100.0	254	2	A44104	cholestenone 5alph
126	21	100.0	255	1	C37165	transcription init
127	21	100.0	255	1	A28567	transcription init

128	21	100.0	256	2	AC2867	hypothetical prote
129	21	100.0	256	2	H97643	arginine/ornithine
130	21	100.0	256	2	H86706	hypothetical prote
131	21	100.0	257	2	F72067	diaminopimelate ep
132	21	100.0	257	2	C86555	diaminopimelate ep
133	21	100.0	257	2	F72116	conserved hypothet
134	21	100.0	257	2	E86507	hypothetical prote
135	21	100.0	258	2	JQ1666	transcription init
136	21	100.0	258	2	B49904	probable regulator
137	21	100.0	258	2	B91189	transcription regu
138	21	100.0	258	2	C86036	transcription regu
139	21	100.0	258	2	AI0975	probable L-lactate
140	21	100.0	259	1	E69315	probable 3-isoprop
141	21	100.0	259	2	G75365	GidA-related prote
142	21	100.0	259	2	F97084	uncharacterized pr
143	21	100.0	260	2	G69758	hypothetical prote
144	21	100.0	262	2	AE3287	thiamine transport
145	21	100.0	269	2	I54432	MHC class II histo
146	21	100.0	269	2	AC0973	lipopolysaccharide
147	21	100.0	270	2	B81199	nitrilase NMB0441
148	21	100.0	270	2	E81834	conserved hypothet
149	21	100.0	270	2	D83072	conserved hypothet
150	21	100.0	271	2	AC2056	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

S59512

probable integrin alpha3 betal - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-May-1999

C;Accession: S59512

R;McCormick, J.I.; Johnstone, R.M.

Biochem. J. 311, 743-751, 1995

A;Title: Identification of the integrin alpha(3)beta(1) as a component of a partially purified A-system amino acid transporter from Ehrlich cell plasma membranes.

A;Reference number: S59512; MUID:96067588; PMID:7487928

A;Accession: S59512

A;Status: preliminary

A;Molecule type: protein

A;Residues: 1-17 <MCC>

Query Match 100.0%; Score 21; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 16;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVRF 4  
 ||||  
 Db 4 DVRF 7

##### RESULT 2

H82806

hypothetical protein XF0431 [imported] - Xylella fastidiosa (strain 9a5c)

C;Species: *Xylella fastidiosa*  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
 C;Accession: H82806  
 R;anonymous, The *Xylella fastidiosa* Consortium of the Organization for  
 Nucleotide Sequencing and Analysis, Sao Paulo, Brazil.  
 Nature 406, 151-157, 2000  
 A;Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
 A;Reference number: A82515; MUID:20365717; PMID:10910347  
 A;Note: for a complete list of authors see reference number A59328 below  
 A;Accession: H82806  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-56 <SIM>  
 A;Cross-references: GB:AE003894; GB:AE003849; NID:g9105267; PIDN:AAF83241.1;  
 GSPDB:GN00128; XFSC:XF0431  
 A;Experimental source: strain 9a5c  
 R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.;  
 Alvarenga, R.; Alves, L.M.C.; Araya, J.E.; Baia, G.S.; Baptista, C.S.; Barros,  
 M.H.; Bonaccorsi, E.D.; Bordin, S.; Bove, J.M.; Briones, M.R.S.; Bueno, M.R.P.;  
 Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.; Colauto, N.B.;  
 Colombo, C.; Costa, F.F.; Costa, M.C.R.; Costa-Neto, C.M.; Coutinho, L.L.;  
 Cristofani, M.; Dias-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.;  
 Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco,  
 M.C.; Frohme, M.; Furlan, L.R.; Garnier, M.; Goldman, G.H.; Goldman, M.H.S.;  
 Gomes, S.L.; Gruber, A.; Ho, P.L.; Hoheisel, J.D.; Junqueira, M.L.; Kemper,  
 E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigret, F.; Lambais, M.R.;  
 Leite, L.C.C.; Lemos, E.G.M.; Lemos, M.V.F.; Lopes, S.A.; Lopes, C.R.; Machado,  
 J.A.; Machado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques,  
 M.V.; Martins, E.A.L.  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.;  
 Miyaki, C.Y.; Monteiro-Vitorello, C.B.; Moon, D.H.; Nagai, M.A.; Nascimento,  
 A.L.T.O.; Netto, L.E.S.; Nhani Jr., A.; Nobrega, F.G.; Nunes, L.R.; Oliveira,  
 M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Paris, A.; Peixoto,  
 B.R.; Pereira, G.A.G.; Pereira Jr., H.A.; Pesquero, J.B.; Quaggio, R.B.;  
 Roberto, P.G.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.;  
 Santelli, R.V.; Sawasaki, H.E.  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da  
 Silveira, J.F.; Silvestri, M.L.Z.; Siqueira, W.J.; de Souza, A.A.; de Souza,  
 A.P.; Terenzi, M.F.; Truffi, D.; Tsai, S.M.; Tsuhako, M.H.; Vallada, H.; Van  
 Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Zago, M.A.; Zatz, M.;  
 Meidanis, J.; Setubal, J.C.  
 A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: XF0431

Query Match 100.0%; Score 21; DB 2; Length 56;  
 Best Local Similarity 100.0%; Pred. No. 59;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVRF 4  
 ||||  
 Db 31 DVRF 34

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:58:36 ; Search time 2.72269 Seconds  
 (without alignments)  
 76.498 Million cell updates/sec

Title: US-10-030-735-54  
 Perfect score: 21  
 Sequence: 1 DVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	21	100.0	24	1	CH60_HELVI	P26317 heliothis v
2	21	100.0	47	1	CH60_LEITA	P56281 leishmania
3	21	100.0	92	1	CF1A_MYCTU	Q10646 mycobacteri
4	21	100.0	99	1	GP45_BPSP1	O48399 bacteriopha
5	21	100.0	102	1	RS24_HALMA	P19953 haloarcula
6	21	100.0	103	1	RL18_PIG	Q95342 sus scrofa
7	21	100.0	103	1	Y273_METJA	Q57721 methanococc
8	21	100.0	110	1	YHBJ_ACTAC	P96769 actinobacil
9	21	100.0	121	1	NQO7_PARDE	P29919 paracoccus
10	21	100.0	121	1	NU3M_PORPU	O99978 porphyra pu
11	21	100.0	121	1	NUA1_RHIME	O68852 rhizobium m
12	21	100.0	122	1	NU3M_RECAM	O21273 reclinomona
13	21	100.0	126	1	NUOA_RHOCA	O84969 rhodobacter
14	21	100.0	127	1	TRX3_YEAST	P25372 saccharomyc
15	21	100.0	135	1	RBFA_RHIME	Q92sw3 rhizobium m
16	21	100.0	141	1	LSHB_HUMAN	P01229 homo sapien
17	21	100.0	144	1	RBFA_BRAJA	Q89wb0 bradyrhizob

18	21	100.0	146	1	MIOC_HAEIN	P44813	haemophilus
19	21	100.0	149	1	RL9_HAEIN	P44349	haemophilus
20	21	100.0	150	1	RECX_LEGPN	P37864	legionella
21	21	100.0	150	1	SSB_THETN	Q8r6m2	thermoanaer
22	21	100.0	152	1	LUXS_HELPU	Q9zwmw8	helicobacte
23	21	100.0	153	1	YCT3_BACPF	Q04445	bacillus ps
24	21	100.0	155	1	LUXS_HELPU	O24931	helicobacte
25	21	100.0	155	1	LUXS_LISIN	Q92c64	listeria in
26	21	100.0	155	1	LUXS_LISMO	Q8y7i9	listeria mo
27	21	100.0	157	1	NIFU_AQUAE	O67045	aquifex aeo
28	21	100.0	158	1	LUXS_LACLA	Q9ciu0	lactococcus
29	21	100.0	161	1	YP49_RALSO	Q8xwc5	ralstonia s
30	21	100.0	162	1	RECX_XANCP	Q8p9x1	xanthomonas
31	21	100.0	165	1	CGHB_HUMAN	P01233	homo sapien
32	21	100.0	167	1	RIMM_STAAM	Q99un1	staphylococ
33	21	100.0	169	1	Y387_LISIN	Q92es1	listeria in
34	21	100.0	172	1	YFIR_ECOLI	P76597	escherichia
35	21	100.0	176	1	Y069_TREPA	O83108	treponema p
36	21	100.0	179	1	Y825_PASMU	Q9cmj9	pasteurella
37	21	100.0	179	1	Y969_VIBPA	Q87r29	vibrio para
38	21	100.0	179	1	YI92_VIBCH	Q9kqv6	vibrio chol
39	21	100.0	180	1	ORN_SCHPO	O94626	schizosacch
40	21	100.0	180	1	YG16_YERPE	Q8zfs2	yersinia pe
41	21	100.0	180	1	YK72_VIBVU	P59274	vibrio vuln
42	21	100.0	180	1	YN69_VIBVY	Q7miz4	vibrio vuln
43	21	100.0	191	1	Y064_TREPA	O83103	treponema p
44	21	100.0	199	1	EQST_ACTEQ	P81439	actinia equ
45	21	100.0	203	1	RECR_MYCLE	O69520	mycobacteri
46	21	100.0	204	1	YWFB_BACSU	P39638	bacillus su
47	21	100.0	207	1	NADD_BACHD	Q9kd91	bacillus ha
48	21	100.0	210	1	URK_DEIRA	Q9rxz5	deinococcus
49	21	100.0	211	1	LIPB_BUCAI	P57356	buchnera ap
50	21	100.0	215	1	THYX_SYNY3	P73053	synechocyst
51	21	100.0	219	1	RR3_ASTLO	P58133	astasia lon
52	21	100.0	221	1	YEK8_YEAST	P40014	saccharomyc
53	21	100.0	222	1	PIMT_CAUCR	Q9a6t6	caulobacter
54	21	100.0	223	1	KCY_BORBR	Q9rnd6	bordetella
55	21	100.0	226	1	KCY_VIBPA	Q87n44	vibrio para
56	21	100.0	226	1	KCY_VIBVU	Q8d8j1	vibrio vuln
57	21	100.0	227	1	KCY_ECOLI	P23863	escherichia
58	21	100.0	227	1	KCY_SALTI	Q8z803	salmonella
59	21	100.0	227	1	KCY_SALTY	Q8zqc4	salmonella
60	21	100.0	229	1	SNG3_HUMAN	O43761	homo sapien
61	21	100.0	230	1	KCY_RALSO	Q8y0y5	ralstonia s
62	21	100.0	230	1	KCY_YERPE	Q8zgb3	yersinia pe
63	21	100.0	233	1	MTGA_NEIGO	Q51005	neisseria g
64	21	100.0	235	1	RSUA_YERPE	Q8zgm2	yersinia pe
65	21	100.0	236	1	RPIA_ANASP	Q8yyg2	anabaena sp
66	21	100.0	236	1	RPIA_SYNEL	Q8djf2	synechococc
67	21	100.0	238	1	RISA_YEAST	P38145	saccharomyc
68	21	100.0	241	1	Y513_METJA	Q57933	methanococc
69	21	100.0	244	1	YE04_YEAST	P32642	saccharomyc
70	21	100.0	246	1	PYRK_CLOPE	Q8xl63	clostridium
71	21	100.0	248	1	PHAB_ACISP	P50203	acinetobact
72	21	100.0	251	1	YY27_MYCTU	Q50701	mycobacteri
73	21	100.0	252	1	COMA_BACSU	O06739	bacillus su
74	21	100.0	254	1	S5A2_RAT	P31214	rattus norv

75	21	100.0	255	1	RPSF_BACLI	P26764	bacillus li
76	21	100.0	255	1	RPSF_BACSU	P07860	bacillus su
77	21	100.0	257	1	DAPF_CHLPN	Q9z833	chlamydia p
78	21	100.0	257	1	YJ85_XYLFT	Q87a46	xylella fas
79	21	100.0	257	1	YQ09_XYLFA	Q9paa9	xylella fas
80	21	100.0	258	1	LLDR_ECOLI	P33233	escherichia
81	21	100.0	258	1	TBP_ACACA	P26354	acanthamoeb
82	21	100.0	271	1	URED_BACSB	Q07400	bacillus sp
83	21	100.0	272	1	Y601_SYNY3	P55175	synechocyst
84	21	100.0	275	1	YAIM_HAEIN	P44556	haemophilus
85	21	100.0	277	1	HEMK_ECOLI	P37186	escherichia
86	21	100.0	277	1	HEMK_SALTY	P40816	salmonella
87	21	100.0	277	1	RIPA_LUFCY	Q00465	luffa cylin
88	21	100.0	280	1	DAPF_ARCFU	O29511	archaeoglob
89	21	100.0	282	1	YPTO_SHEVI	Q9s0k9	shewanella
90	21	100.0	283	1	RFAZ_ECOLI	P27241	escherichia
91	21	100.0	284	1	Y3G4_SHEON	Q8eae2	shewanella
92	21	100.0	284	1	YHBJ_ECOLI	P33995	escherichia
93	21	100.0	284	1	YHBJ_SALTI	Q8z3g1	salmonella
94	21	100.0	284	1	YHBJ_SALTY	Q8zlr8	salmonella
95	21	100.0	284	1	YI34_THETN	Q8r8z8	thermoanaer
96	21	100.0	284	1	YPTO_KLEOX	P17163	klebsiella
97	21	100.0	284	1	YPTO_PROMI	Q9za87	proteus mir
98	21	100.0	284	1	YZ86_YERPE	Q8zb41	yersinia pe
99	21	100.0	285	1	YB46_HAEIN	P45071	haemophilus
100	21	100.0	285	1	YX47_STRCO	Q9x8h0	streptomyce
101	21	100.0	286	1	HEMK_VIBCH	Q9kq26	vibrio chol
102	21	100.0	286	1	YP32_VIBCH	Q9kp47	vibrio chol
103	21	100.0	287	1	Y584_HAEDU	Q917v3	haemophilus
104	21	100.0	287	1	Y695_VIBVU	Q8dea0	vibrio vuln
105	21	100.0	287	1	YPTO_VIBHA	Q9kiq6	vibrio harv
106	21	100.0	287	1	YQ73_VIBPA	Q87ld9	vibrio para
107	21	100.0	288	1	Y169_PASMU	Q9cp85	pasteurella
108	21	100.0	288	1	Y705_BIFLO	Q8g6d8	bifidobacte
109	21	100.0	290	1	VG33_HSVI1	Q00118	ictalurid h
110	21	100.0	290	1	YA89_FUSNN	Q8rel1	fusobacteri
111	21	100.0	291	1	YO74_LISMO	Q8y4g9	listeria mo
112	21	100.0	291	1	YQ17_LISIN	Q928b9	listeria in
113	21	100.0	292	1	DHYS_PYRAE	Q8zt09	pyrobaculum
114	21	100.0	292	1	HEMK_HAEIN	P45253	haemophilus
115	21	100.0	293	1	YO68_OCEIH	Q8enl3	oceanobacil
116	21	100.0	294	1	MDH_ARCFU	O08349	archaeoglob
117	21	100.0	294	1	Y354_CLOPE	Q8xnh9	clostridium
118	21	100.0	294	1	Y511_CLOAB	Q97lp3	clostridium
119	21	100.0	294	1	Y779_LACPL	Q88yi4	lactobacill
120	21	100.0	295	1	DAPF_METJA	Q58519	methanococc
121	21	100.0	295	1	YHYA_PSESN	Q01265	pseudomonas
122	21	100.0	295	1	YJIE_LACLA	Q9cgy1	lactococcus
123	21	100.0	295	1	YVCJ_BACSU	O06973	bacillus su
124	21	100.0	295	1	YZ69_BACHD	Q9k705	bacillus ha
125	21	100.0	296	1	Y576_STRA3	Q8e6i7	streptococc
126	21	100.0	296	1	Y652_STRPY	Q9a0r8	streptococc
127	21	100.0	296	1	Y713_STRP8	Q8p1t8	streptococc
128	21	100.0	296	1	YD06_STRMU	Q8dtm7	streptococc
129	21	100.0	296	1	YE24_STRR6	Q8dp10	streptococc
130	21	100.0	296	1	YF66_STRPN	Q97pn7	streptococc
131	21	100.0	298	1	HEMK_PASMU	Q9cn82	pasteurella



132	21	100.0	298	1	Y563_MYCLE	Q9ccp0 mycobacteri
133	21	100.0	300	1	FTRL_METKA	Q8tx60 methanopyru
134	21	100.0	301	1	YE21_MYCTU	P71690 mycobacteri
135	21	100.0	302	1	Y548_STAEP	Q8cte3 staphylococ
136	21	100.0	303	1	Y765_STAAM	Q99vl1 staphylococ
137	21	100.0	304	1	K2S2_HUMAN	P43631 homo sapien
138	21	100.0	304	1	PHEA_AMEYE	Q44104 amycolatops
139	21	100.0	307	1	SYGA_XYLFA	Q9pc25 xylella fas
140	21	100.0	307	1	SYGA_XYLFT	Q87d46 xylella fas
141	21	100.0	308	1	SYGA_BRUME	Q8yfy0 brucella me
142	21	100.0	308	1	SYGA_BRUSU	Q8g2c1 brucella su
143	21	100.0	309	1	AURC_HUMAN	Q9uqb9 homo sapien
144	21	100.0	309	1	YF91_CORGL	Q8nq56 corynebacte
145	21	100.0	310	1	SYGA_AGR5	Q8uhp1 agrobacteri
146	21	100.0	314	1	HEMZ_LACLA	Q9cfb4 lactococcus
147	21	100.0	315	1	SYGA_PSESM	Q88b36 pseudomonas
148	21	100.0	317	1	FECR_ECOLI	P23485 escherichia
149	21	100.0	318	1	SYGA_MORCA	P77892 moraxella c
150	21	100.0	325	1	MOAA_EUBAC	Q93kd1 eubacterium

#### ALIGNMENTS

#### RESULT 1

CH60\_HELVI

ID CH60\_HELVI STANDARD; PRT; 24 AA.

AC P26317;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE 60 kDa chaperonin, mitochondrial (P60) (Fragment).

OS *Heliothis virescens* (Noctuid moth) (Owlet moth).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;

OC Noctuidae; *Heliothinae*; *Heliothis*.

OX NCBI\_TaxID=7102;

RN [1]

RP SEQUENCE.

RC TISSUE=Testis;

RX MEDLINE=90339485; PubMed=1974308;

RA Miller S.G., Leclerc R.F., Erdos G.W.;

RT "Identification and characterization of a testis-specific isoform of  
a chaperonin in a moth, *Heliothis virescens*.";

RL J. Mol. Biol. 214:407-422.(1990).

CC -!- FUNCTION: Implicated in mitochondrial protein import and  
macromolecular assembly. May facilitate the correct folding of  
imported proteins. May also prevent misfolding and promote the  
refolding and proper assembly of unfolded polypeptides generated  
under stress conditions in the mitochondrial matrix (By  
similarity).

CC -!- SUBUNIT: Forms a single seven-member ring complex, in tight  
association with the p63 protein.

CC -!- SUBCELLULAR LOCATION: Mitochondrial.

CC -!- TISSUE SPECIFICITY: Testis.

CC -!- DEVELOPMENTAL STAGE: From the second half of the larval final-  
instar, through the first two days of pupal development.

CC -!- MISCELLANEOUS: Shows ATPase activity.  
 CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
 DR InterPro; IPR001844; Chaprnin\_Cpn60.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; PARTIAL.  
 KW Chaperone; ATP-binding; Testis; Mitochondrion.  
 FT NON\_TER 24 24  
 SQ SEQUENCE 24 AA; 2531 MW; 2B34508F8CA981CF CRC64;

Query Match 100.0%; Score 21; DB 1; Length 24;  
 Best Local Similarity 100.0%; Pred. No. 9.7;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 DVRF 4  
 ||||  
 Db 3 DVRF 6

## RESULT 2

CH60\_LEITA  
 ID CH60\_LEITA STANDARD; PRT; 47 AA.  
 AC P56281;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Chaperonin HSP60, mitochondrial precursor (Protein Cpn60) (groEL protein) (Heat shock protein 60) (Fragment).  
 GN HSP60.  
 OS Leishmania tarentolae (Sauroleishmania tarentolae).  
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
 OX NCBI\_TaxID=5689;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 9-18.  
 RX MEDLINE=96360469; PubMed=8719252;  
 RA Bringaud F., Peyruchaud S., Baltz D., Giroud C., Simpson L., Baltz T.;  
 RT "Molecular characterization of the mitochondrial heat shock protein  
 RT 60 gene from Trypanosoma brucei.";  
 RL Mol. Biochem. Parasitol. 74:119-123(1995).  
 CC -!- FUNCTION: Implicated in mitochondrial protein import and  
 CC macromolecular assembly. May facilitate the correct folding of  
 CC imported proteins. May also prevent misfolding and promote the  
 CC refolding and proper assembly of unfolded polypeptides generated  
 CC under stress conditions in the mitochondrial matrix.  
 CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).  
 CC -!- INDUCTION: By heat shock (By similarity).  
 CC -!- SIMILARITY: Belongs to the chaperonin (HSP60) family.  
 DR HSSP; P06139; 1AON.  
 DR InterPro; IPR001844; Chaprnin\_Cpn60.  
 DR InterPro; IPR008950; GroEL-ATPase.  
 DR PROSITE; PS00296; CHAPERONINS\_CPN60; PARTIAL.  
 KW Chaperone; ATP-binding; Transit peptide; Mitochondrion; Heat shock.  
 FT TRANSIT 1 8 MITOCHONDRION.  
 FT CHAIN 9 >47 CHAPERONIN HSP60.  
 FT NON\_TER 47 47  
 SQ SEQUENCE 47 AA; 5025 MW; 92589CD137336008 CRC64;

Query Match 100.0%; Score 21; DB 1; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 20;

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:09:21 ; Search time 16.1345 Seconds  
 (without alignments)  
 78.222 Million cell updates/sec

Title: US-10-030-735-54  
 Perfect score: 21  
 Sequence: 1 DVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SPTREMBL\_25:\*  
 1: sp\_archaea:\*  
 2: sp\_bacteria:\*  
 3: sp\_fungi:\*  
 4: sp\_human:\*  
 5: sp\_invertebrate:\*  
 6: sp\_mammal:\*  
 7: sp\_mhc:\*  
 8: sp\_organelle:\*  
 9: sp\_phage:\*  
 10: sp\_plant:\*  
 11: sp\_rodent:\*  
 12: sp\_virus:\*  
 13: sp\_vertebrate:\*  
 14: sp\_unclassified:\*  
 15: sp\_rvirus:\*  
 16: sp\_bacteriap:\*  
 17: sp\_archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	Score	Match	Length	DB	ID	Description
No.						

---

1	21	100.0	28	4	Q9UCG8	Q9ucg8 homo sapien
2	21	100.0	30	4	Q9UDI0	Q9udi0 homo sapien
3	21	100.0	42	16	Q7UHZ7	Q7uhz7 rhodopirell
4	21	100.0	47	7	Q9GIW8	Q9giw8 monachus sc
5	21	100.0	52	16	Q8CU29	Q8cu29 staphylococ
6	21	100.0	56	16	Q9PG71	Q9pg71 xylella fas
7	21	100.0	57	7	Q30002	Q30002 homo sapien
8	21	100.0	59	9	Q856L4	Q856l4 mycobacteri
9	21	100.0	68	7	Q30121	Q30121 homo sapien
10	21	100.0	72	8	Q9G8W3	Q9g8w3 rhodomonas
11	21	100.0	73	7	Q95381	Q95381 homo sapien
12	21	100.0	73	9	Q857S7	Q857s7 mycobacteri
13	21	100.0	76	16	Q988J8	Q988j8 rhizobium l
14	21	100.0	77	4	Q9P1E2	Q9p1e2 homo sapien
15	21	100.0	77	16	Q82X22	Q82x22 nitrosomona
16	21	100.0	79	11	Q8BZG9	Q8bzg9 mus musculu
17	21	100.0	80	7	Q9TPA8	Q9tpa8 homo sapien
18	21	100.0	81	6	Q9TSK7	Q9tsk7 macaca fasc
19	21	100.0	81	7	Q99947	Q99947 homo sapien
20	21	100.0	81	7	Q30731	Q30731 macaca neme
21	21	100.0	82	2	O69402	O69402 desulfovibr
22	21	100.0	82	7	Q31550	Q31550 metriaclima
23	21	100.0	82	7	Q31453	Q31453 melanochrom
24	21	100.0	82	7	Q31451	Q31451 melanochrom
25	21	100.0	82	7	Q31435	Q31435 melanochrom
26	21	100.0	82	7	Q31539	Q31539 metriaclima
27	21	100.0	82	7	Q31434	Q31434 melanochrom
28	21	100.0	82	7	Q30949	Q30949 pan troglod
29	21	100.0	85	12	Q91KJ4	Q91kj4 hepatitis c
30	21	100.0	86	9	Q855H4	Q855h4 mycobacteri
31	21	100.0	87	7	Q9TP51	Q9tp51 ovis aries
32	21	100.0	87	16	Q9CCW3	Q9ccw3 mycobacteri
33	21	100.0	88	2	Q9ZET6	Q9zet6 xanthobacte
34	21	100.0	89	7	Q9BD02	Q9bd02 homo sapien
35	21	100.0	90	7	Q30100	Q30100 homo sapien
36	21	100.0	90	9	Q855R1	Q855r1 mycobacteri
37	21	100.0	91	16	Q9KXU7	Q9kxu7 streptomyce
38	21	100.0	91	16	Q82GG4	Q82gg4 streptomyce
39	21	100.0	92	2	Q9EUR9	Q9eur9 streptomyce
40	21	100.0	92	16	Q8NMI0	Q8nmi0 corynebacte
41	21	100.0	92	16	Q9S2H0	Q9s2h0 streptomyce
42	21	100.0	94	2	Q9RBP1	Q9rbp1 rhodococcus
43	21	100.0	95	9	Q9B082	Q9b082 mycobacteri
44	21	100.0	95	16	Q9JZC2	Q9jzc2 neisseria m
45	21	100.0	95	16	Q82D24	Q82d24 streptomyce
46	21	100.0	96	16	Q9JMR7	Q9jmr7 escherichia
47	21	100.0	96	16	Q88W58	Q88w58 lactobacill
48	21	100.0	100	9	Q853C3	Q853c3 mycobacteri
49	21	100.0	101	1	Q9P9E8	Q9p9e8 methanococc
50	21	100.0	103	12	Q9E977	Q9e977 hepatitis c
51	21	100.0	105	4	Q9UNB8	Q9unb8 homo sapien
52	21	100.0	105	4	Q9UNB6	Q9unb6 homo sapien
53	21	100.0	106	10	Q9SVR6	Q9svr6 arabidopsis
54	21	100.0	107	16	Q88G11	Q88g11 pseudomonas
55	21	100.0	108	9	Q8W6U2	Q8w6u2 bacterioph
56	21	100.0	108	11	Q8C3G0	Q8c3g0 mus musculu
57	21	100.0	109	16	Q9A615	Q9a615 caulobacter

58	21	100.0	109	16	Q9FBH2	Q9fbh2 streptococc
59	21	100.0	109	16	Q8CZ47	Q8cz47 streptococc
60	21	100.0	110	11	Q9JL83	Q9jl83 mus musculu
61	21	100.0	113	3	Q9Y7S1	Q9y7s1 schizosacch
62	21	100.0	115	17	Q97YS7	Q97ys7 sulfolobus
63	21	100.0	116	2	Q7WW39	Q7ww39 yersinia be
64	21	100.0	118	8	Q02666	Q02666 podospora a
65	21	100.0	118	16	Q988K6	Q988k6 rhizobium l
66	21	100.0	118	16	Q83BQ5	Q83bq5 coxiella bu
67	21	100.0	120	3	Q96W45	Q96w45 ophiostoma
68	21	100.0	120	8	Q8LWY2	Q8lwy2 laminaria d
69	21	100.0	120	8	Q94YZ1	Q94yz1 pylaiella l
70	21	100.0	121	8	Q9MG92	Q9mg92 chrysodidym
71	21	100.0	121	16	Q89KI6	Q89ki6 bradyrhizob
72	21	100.0	122	2	Q7WW11	Q7ww11 yersinia mo
73	21	100.0	123	8	Q9G8Q2	Q9g8q2 naegleria g
74	21	100.0	123	12	Q68895	Q68895 hepatitis c
75	21	100.0	123	17	Q9HL85	Q9hl85 thermoplasm
76	21	100.0	124	16	Q985D3	Q985d3 rhizobium l
77	21	100.0	124	16	Q92GI4	Q92gi4 rickettsia
78	21	100.0	125	16	Q9A6X0	Q9a6x0 caulobacter
79	21	100.0	126	17	Q97XQ2	Q97xq2 sulfolobus
80	21	100.0	127	16	Q8CMB5	Q8cmb5 streptococc
81	21	100.0	128	2	Q9R2S5	Q9r2s5 yersinia en
82	21	100.0	128	2	Q9RIF9	Q9rif9 yersinia en
83	21	100.0	128	12	Q81481	Q81481 hepatitis c
84	21	100.0	130	2	P70743	P70743 acidimicrob
85	21	100.0	130	8	Q9G8Z8	Q9g8z8 ochromonas
86	21	100.0	130	10	Q8GWW5	Q8gww5 arabidopsis
87	21	100.0	132	2	Q9RAY7	Q9ray7 frankia sp.
88	21	100.0	132	16	Q9RT34	Q9rt34 deinococcus
89	21	100.0	132	16	Q9CBV0	Q9cbv0 mycobacteri
90	21	100.0	134	13	P70072	P70072 gallus gall
91	21	100.0	135	2	Q8GCH4	Q8gch4 rhodovulum
92	21	100.0	136	4	Q8WXL0	Q8wxl0 homo sapien
93	21	100.0	136	6	Q8WNB8	Q8wnb8 pongo pygma
94	21	100.0	136	16	Q8YZF9	Q8yzf9 anabaena sp
95	21	100.0	138	2	Q7WW37	Q7ww37 yersinia en
96	21	100.0	138	12	Q84590	Q84590 paramecium
97	21	100.0	139	16	Q82NK0	Q82nk0 streptomyce
98	21	100.0	140	2	Q52903	Q52903 rhizobium m
99	21	100.0	141	16	Q8DI47	Q8di47 synechococc
100	21	100.0	141	16	Q88KM2	Q88km2 pseudomonas
101	21	100.0	141	16	Q831H9	Q831h9 enterococcu
102	21	100.0	142	2	Q7WW41	Q7ww41 yersinia be
103	21	100.0	142	6	Q9MZV3	Q9mzv3 canis famil
104	21	100.0	142	13	Q801X7	Q801x7 anguilla ja
105	21	100.0	143	2	Q93A93	Q93a93 nitrosococc
106	21	100.0	144	2	Q84HC7	Q84hc7 streptomyce
107	21	100.0	144	2	Q7WW32	Q7ww32 yersinia fr
108	21	100.0	144	16	Q7WF68	Q7wf68 bordetella
109	21	100.0	144	16	Q7W3T8	Q7w3t8 bordetella
110	21	100.0	144	16	Q7W053	Q7w053 bordetella
111	21	100.0	144	16	Q7USR9	Q7usr9 rhodopirell
112	21	100.0	145	4	Q96L30	Q96l30 homo sapien
113	21	100.0	145	6	Q29507	Q29507 oryctolagus
114	21	100.0	146	2	Q7WW22	Q7ww22 yersinia kr

115	21	100.0	148	2	Q7WW21	Q7ww21 yersinia kr
116	21	100.0	148	10	Q9LXH1	Q9lxlh1 arabidopsis
117	21	100.0	149	11	Q925S6	Q925s6 mus musculu
118	21	100.0	149	16	Q7V7F9	Q7v7f9 prochloroco
119	21	100.0	149	16	Q7U731	Q7u731 synechococc
120	21	100.0	150	2	O87454	O87454 vibrio chol
121	21	100.0	150	2	Q7WW27	Q7ww27 yersinia in
122	21	100.0	150	10	Q84US1	Q84us1 oryza sativ
123	21	100.0	150	16	Q8R6M2	Q8r6m2 thermoanaer
124	21	100.0	150	16	Q8PE55	Q8pe55 xanthomonas
125	21	100.0	151	2	Q7WW44	Q7ww44 yersinia be
126	21	100.0	151	2	Q7WW36	Q7ww36 yersinia en
127	21	100.0	151	2	Q7WW14	Q7ww14 yersinia mo
128	21	100.0	151	16	Q88EX9	Q88ex9 pseudomonas
129	21	100.0	152	2	Q7WW35	Q7ww35 yersinia en
130	21	100.0	152	16	O88044	O88044 streptomyce
131	21	100.0	153	2	Q7WW33	Q7ww33 yersinia en
132	21	100.0	153	4	Q9UNB5	Q9unb5 homo sapien
133	21	100.0	153	17	Q8TU50	Q8tu50 methanosarc
134	21	100.0	154	2	Q52339	Q52339 unidentifie
135	21	100.0	154	5	Q17197	Q17197 brugia mala
136	21	100.0	155	2	Q9WWJ5	Q9wwj5 burkholderi
137	21	100.0	156	16	Q828B4	Q828b4 streptomyce
138	21	100.0	157	6	Q8WNC4	Q8wnc4 pongo pygma
139	21	100.0	157	6	Q8WNC5	Q8wnc5 pongo pygma
140	21	100.0	157	6	Q8WNC7	Q8wnc7 pongo pygma
141	21	100.0	157	6	Q8WNC6	Q8wnc6 pongo pygma
142	21	100.0	157	16	Q87Z07	Q87z07 pseudomonas
143	21	100.0	158	2	Q54354	Q54354 streptomyce
144	21	100.0	158	2	Q93UP5	Q93up5 synechococc
145	21	100.0	158	2	Q9X3A8	Q9x3a8 variovorax
146	21	100.0	159	2	Q7WW15	Q7ww15 yersinia mo
147	21	100.0	159	2	Q7WW12	Q7ww12 yersinia mo
148	21	100.0	159	4	Q8WXL1	Q8wxl1 homo sapien
149	21	100.0	160	2	Q7WW09	Q7ww09 yersinia mo
150	21	100.0	160	16	Q89LL2	Q89ll2 bradyrhizob

#### ALIGNMENTS

##### RESULT 1

Q9UCG8

ID Q9UCG8 PRELIMINARY; PRT; 28 AA.

AC Q9UCG8;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Urinary gonadotrophin peptide (Fragment).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE.

RX MEDLINE=93229246; PubMed=8471426;

RA Kardana A., Bagshawe K.D., Coles B., Read D., Taylor M.;

RT "Characterisation of UGP and its relationship with beta-core  
RT fragment."  
RL Br. J. Cancer 67:686-692(1993).  
DR HSSP; P01233; 1HCN.  
DR InterPro; IPR006208; Cys\_knot.  
DR Pfam; PF00007; Cys\_knot; 1.  
SQ SEQUENCE 28 AA; 3150 MW; EF76A484C83D9071 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 28;  
Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVRF 4  
|||  
Db 7 DVRF 10

## RESULT 2

Q9UD10

ID Q9UD10 PRELIMINARY; PRT; 30 AA.  
AC Q9UD10;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HLH beta core fragment, luteinizing hormone-beta core fragment  
DE (Fragments).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=93373892; PubMed=7689962;  
RA Birken S., Chen Y., Gawinowicz M.A., Agosto G.M., Canfield R.E.,  
RA Hartree A.S.;  
RT "Structure and significance of human luteinizing hormone-beta core  
RT fragment purified from human pituitary extracts."  
RL Endocrinology 133:985-989(1993).  
FT NON\_TER 1 1  
FT NON\_CONS 15 16  
FT NON\_TER 30 30  
SQ SEQUENCE 30 AA; 3585 MW; 58A2308EDD4AE663 CRC64;

Query Match 100.0%; Score 21; DB 4; Length 30;  
Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DVRF 4  
|||  
Db 22 DVRF 25

## RESULT 3

Q7UHZ7

ID Q7UHZ7 PRELIMINARY; PRT; 42 AA.  
AC Q7UHZ7;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:16:36 ; Search time 6.28571 Seconds  
 (without alignments)  
 32.853 Million cell updates/sec

Title: US-10-030-735-55  
 Perfect score: 21  
 Sequence: 1 NLRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Issued\_Patents\_AA:\*  
 1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
 2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
 3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
 4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
 5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	21	100.0	12	1 US-08-511-662-8	Sequence 8, Appli
2	21	100.0	12	4 US-09-459-749D-6	Sequence 6, Appli
3	21	100.0	12	5 PCT-US96-12632-8	Sequence 8, Appli
4	21	100.0	19	2 US-09-025-706-11	Sequence 11, Appl
5	21	100.0	19	3 US-09-025-622-11	Sequence 11, Appl
6	21	100.0	29	3 US-08-933-100B-7	Sequence 7, Appli
7	21	100.0	62	4 US-09-328-352-5258	Sequence 5258, Ap
8	21	100.0	63	4 US-09-489-039A-14108	Sequence 14108, A
9	21	100.0	66	4 US-09-732-210-823	Sequence 823, App
10	21	100.0	66	4 US-09-732-210-824	Sequence 824, App
11	21	100.0	66	4 US-09-673-395A-485	Sequence 485, App



12	21	100.0	70	4	US-09-489-039A-14229	Sequence 14229, A
13	21	100.0	71	4	US-09-134-000C-5815	Sequence 5815, Ap
14	21	100.0	75	4	US-09-732-210-833	Sequence 833, App
15	21	100.0	76	4	US-09-543-681A-5143	Sequence 5143, Ap
16	21	100.0	78	4	US-09-134-001C-3279	Sequence 3279, Ap
17	21	100.0	78	4	US-09-543-681A-5624	Sequence 5624, Ap
18	21	100.0	85	2	US-08-480-229C-5	Sequence 5, Appli
19	21	100.0	85	2	US-08-659-235C-5	Sequence 5, Appli
20	21	100.0	101	4	US-09-540-236-2806	Sequence 2806, Ap
21	21	100.0	109	4	US-09-325-932A-144	Sequence 144, App
22	21	100.0	115	4	US-09-540-236-2698	Sequence 2698, Ap
23	21	100.0	117	4	US-09-540-236-2827	Sequence 2827, Ap
24	21	100.0	118	4	US-09-489-039A-9348	Sequence 9348, Ap
25	21	100.0	119	4	US-08-890-865A-17	Sequence 17, Appl
26	21	100.0	123	2	US-08-588-258B-1	Sequence 1, Appli
27	21	100.0	123	2	US-08-588-258B-39	Sequence 39, Appl
28	21	100.0	123	3	US-08-460-505-1	Sequence 1, Appli
29	21	100.0	123	5	PCT-US96-08295-1	Sequence 1, Appli
30	21	100.0	123	5	PCT-US96-08295-39	Sequence 39, Appl
31	21	100.0	124	1	US-08-276-852-127	Sequence 127, App
32	21	100.0	124	1	US-08-276-852-132	Sequence 132, App
33	21	100.0	124	1	US-08-899-575-127	Sequence 127, App
34	21	100.0	124	1	US-08-899-575-132	Sequence 132, App
35	21	100.0	124	1	US-08-899-575-127	Sequence 127, App
36	21	100.0	124	1	US-08-899-575-132	Sequence 132, App
37	21	100.0	124	5	PCT-US95-08743-127	Sequence 127, App
38	21	100.0	124	5	PCT-US95-08743-132	Sequence 132, App
39	21	100.0	125	1	US-08-276-852-124	Sequence 124, App
40	21	100.0	125	1	US-08-276-852-128	Sequence 128, App
41	21	100.0	125	1	US-08-899-575-124	Sequence 124, App
42	21	100.0	125	1	US-08-899-575-128	Sequence 128, App
43	21	100.0	125	1	US-08-899-575-124	Sequence 124, App
44	21	100.0	125	1	US-08-899-575-128	Sequence 128, App
45	21	100.0	125	5	PCT-US95-08743-124	Sequence 124, App
46	21	100.0	125	5	PCT-US95-08743-128	Sequence 128, App
47	21	100.0	126	1	US-08-276-852-123	Sequence 123, App
48	21	100.0	126	1	US-08-899-575-123	Sequence 123, App
49	21	100.0	126	1	US-08-899-575-123	Sequence 123, App
50	21	100.0	126	5	PCT-US95-08743-123	Sequence 123, App
51	21	100.0	129	4	US-09-370-838-214	Sequence 214, App
52	21	100.0	130	4	US-09-328-352-5141	Sequence 5141, Ap
53	21	100.0	153	4	US-09-134-000C-3493	Sequence 3493, Ap
54	21	100.0	156	2	US-08-162-402B-18	Sequence 18, Appl
55	21	100.0	157	1	US-08-257-999-4	Sequence 4, Appli
56	21	100.0	166	4	US-09-328-352-5740	Sequence 5740, Ap
57	21	100.0	174	3	US-09-153-586-23	Sequence 23, Appl
58	21	100.0	179	3	US-09-042-771-4	Sequence 4, Appli
59	21	100.0	184	1	US-08-353-550-2	Sequence 2, Appli
60	21	100.0	184	2	US-08-551-687-2	Sequence 2, Appli
61	21	100.0	186	4	US-09-543-681A-6955	Sequence 6955, Ap
62	21	100.0	189	3	US-08-821-994-79	Sequence 79, Appl
63	21	100.0	195	4	US-09-489-039A-8601	Sequence 8601, Ap
64	21	100.0	197	4	US-09-252-991A-23101	Sequence 23101, A
65	21	100.0	198	4	US-09-489-039A-8826	Sequence 8826, Ap
66	21	100.0	198	4	US-09-540-236-2237	Sequence 2237, Ap
67	21	100.0	199	4	US-09-252-991A-26198	Sequence 26198, A
68	21	100.0	203	4	US-09-198-452A-196	Sequence 196, App

69	21	100.0	208	4	US-09-134-001C-3478	Sequence 3478, Ap
70	21	100.0	213	4	US-09-489-039A-7236	Sequence 7236, Ap
71	21	100.0	219	4	US-09-540-236-3066	Sequence 3066, Ap
72	21	100.0	225	3	US-08-821-994-49	Sequence 49, Appl
73	21	100.0	225	3	US-08-821-994-51	Sequence 51, Appl
74	21	100.0	230	3	US-08-821-994-50	Sequence 50, Appl
75	21	100.0	232	4	US-09-328-352-8189	Sequence 8189, Ap
76	21	100.0	237	4	US-09-134-000C-5226	Sequence 5226, Ap
77	21	100.0	238	3	US-09-277-716-32	Sequence 32, Appl
78	21	100.0	238	4	US-09-609-161B-32	Sequence 32, Appl
79	21	100.0	241	4	US-09-252-991A-25199	Sequence 25199, A
80	21	100.0	244	3	US-08-821-994-80	Sequence 80, Appl
81	21	100.0	254	2	US-08-207-481-20	Sequence 20, Appl
82	21	100.0	254	5	PCT-US95-02689-20	Sequence 20, Appl
83	21	100.0	258	4	US-09-540-236-2125	Sequence 2125, Ap
84	21	100.0	259	3	US-09-185-160-9	Sequence 9, Appli
85	21	100.0	261	2	US-08-353-476-69	Sequence 69, Appl
86	21	100.0	262	2	US-08-353-476-70	Sequence 70, Appl
87	21	100.0	262	4	US-09-252-991A-30480	Sequence 30480, A
88	21	100.0	264	2	US-08-484-905-120	Sequence 120, App
89	21	100.0	264	3	US-08-481-985B-120	Sequence 120, App
90	21	100.0	264	3	US-08-370-476-120	Sequence 120, App
91	21	100.0	264	4	US-09-399-081A-10	Sequence 10, Appl
92	21	100.0	271	3	US-08-536-891A-1	Sequence 1, Appli
93	21	100.0	271	3	US-09-181-183-36	Sequence 36, Appl
94	21	100.0	271	3	US-08-933-100B-14	Sequence 14, Appl
95	21	100.0	271	4	US-09-280-040-36	Sequence 36, Appl
96	21	100.0	271	4	US-09-277-700-36	Sequence 36, Appl
97	21	100.0	271	4	US-09-366-009-1	Sequence 1, Appli
98	21	100.0	271	4	US-08-809-156B-1	Sequence 1, Appli
99	21	100.0	271	4	US-09-043-981-1	Sequence 1, Appli
100	21	100.0	271	4	US-09-874-585D-36	Sequence 36, Appl
101	21	100.0	271	4	US-09-874-585D-53	Sequence 53, Appl
102	21	100.0	279	3	US-09-042-771-2	Sequence 2, Appli
103	21	100.0	280	4	US-09-489-039A-9259	Sequence 9259, Ap
104	21	100.0	284	4	US-09-543-681A-7910	Sequence 7910, Ap
105	21	100.0	285	4	US-09-107-532A-3967	Sequence 3967, Ap
106	21	100.0	290	4	US-09-543-681A-7488	Sequence 7488, Ap
107	21	100.0	296	2	US-08-836-854-4	Sequence 4, Appli
108	21	100.0	304	4	US-09-134-000C-5217	Sequence 5217, Ap
109	21	100.0	313	4	US-09-328-352-7010	Sequence 7010, Ap
110	21	100.0	318	2	US-08-336-198C-3	Sequence 3, Appli
111	21	100.0	318	4	US-09-252-991A-17173	Sequence 17173, A
112	21	100.0	318	4	US-09-252-991A-32357	Sequence 32357, A
113	21	100.0	318	4	US-09-184-965-3	Sequence 3, Appli
114	21	100.0	319	4	US-09-543-681A-5176	Sequence 5176, Ap
115	21	100.0	324	2	US-08-671-320-11	Sequence 11, Appl
116	21	100.0	324	2	US-08-868-577-11	Sequence 11, Appl
117	21	100.0	324	4	US-09-207-914-11	Sequence 11, Appl
118	21	100.0	326	4	US-09-134-000C-6054	Sequence 6054, Ap
119	21	100.0	328	3	US-08-821-994-84	Sequence 84, Appl
120	21	100.0	334	4	US-09-198-452A-923	Sequence 923, App
121	21	100.0	340	3	US-09-120-365-60	Sequence 60, Appl
122	21	100.0	340	3	US-09-515-039-60	Sequence 60, Appl
123	21	100.0	346	4	US-09-489-039A-8022	Sequence 8022, Ap
124	21	100.0	351	1	US-09-500-651-2	Sequence 2, Appli
125	21	100.0	351	2	US-08-813-591-2	Sequence 2, Appli

126	21	100.0	355	4	US-09-634-238-384	Sequence 384, App
127	21	100.0	363	4	US-09-252-991A-19920	Sequence 19920, A
128	21	100.0	367	4	US-09-252-991A-20488	Sequence 20488, A
129	21	100.0	368	2	US-08-836-854-17	Sequence 17, Appl
130	21	100.0	369	1	US-08-181-629A-4	Sequence 4, Appli
131	21	100.0	372	4	US-09-107-532A-4779	Sequence 4779, Ap
132	21	100.0	373	4	US-09-252-991A-31094	Sequence 31094, A
133	21	100.0	374	3	US-08-821-994-67	Sequence 67, Appl
134	21	100.0	374	3	US-08-821-994-68	Sequence 68, Appl
135	21	100.0	374	3	US-08-821-994-69	Sequence 69, Appl
136	21	100.0	374	3	US-08-821-994-70	Sequence 70, Appl
137	21	100.0	383	4	US-09-459-749D-17	Sequence 17, Appl
138	21	100.0	386	4	US-09-325-932A-143	Sequence 143, App
139	21	100.0	387	4	US-09-328-352-5367	Sequence 5367, Ap
140	21	100.0	389	4	US-09-464-035A-7	Sequence 7, Appli
141	21	100.0	392	3	US-08-689-421-19	Sequence 19, Appl
142	21	100.0	392	3	US-09-389-528-19	Sequence 19, Appl
143	21	100.0	392	3	US-09-181-827A-19	Sequence 19, Appl
144	21	100.0	403	3	US-08-776-246-4	Sequence 4, Appli
145	21	100.0	403	4	US-09-489-039A-12412	Sequence 12412, A
146	21	100.0	404	4	US-09-198-452A-718	Sequence 718, App
147	21	100.0	408	4	US-09-543-681A-5372	Sequence 5372, Ap
148	21	100.0	409	4	US-09-543-681A-4445	Sequence 4445, Ap
149	21	100.0	411	4	US-09-252-991A-18602	Sequence 18602, A
150	21	100.0	412	4	US-09-423-439-10	Sequence 10, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-511-662-8

; Sequence 8, Application US/08511662

; Patent No. 5807552

; GENERAL INFORMATION:

; APPLICANT: Stanton, G. John

; APPLICANT: Hughes, Jr., Thomas K.

; APPLICANT: Smith, Eric M.

; TITLE OF INVENTION: Compositions for Conferring Immunogenicity

; TITLE OF INVENTION: to a Substance and Uses Thereof

; NUMBER OF SEQUENCES: 12

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Arnold, White & Durkee

; STREET: P.O. Box 4433

; CITY: Houston

; STATE: TX

; COUNTRY: USA

; ZIP: 77210

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/511,662

; FILING DATE: Concurrently herewith

; CLASSIFICATION: 530

```

; ATTORNEY/AGENT INFORMATION:
;   NAME: Hodgins, Daniel S.
;   REGISTRATION NUMBER: 31,026
;   REFERENCE/DOCKET NUMBER: UTSG:162/HOD
; TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 515/418-3000
;   TELEFAX: 512/474-7577
;   TELEX: NA
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 12 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: peptide
US-08-511-662-8

```

```

Query Match          100.0%; Score 21; DB 1; Length 12;
Best Local Similarity 100.0%; Pred. No. 26;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 NLR 4
        |||
Db      7 NLR 10

```

```

RESULT 2
US-09-459-749D-6
; Sequence 6, Application US/09459749D
; Patent No. 6464975
; GENERAL INFORMATION:
; APPLICANT: Millis, Albert J. T.
; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
; FILE REFERENCE: 0794.016A
; CURRENT APPLICATION NUMBER: US/09/459,749D
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,856
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:SITE
;   NAME/KEY: SITE
;   LOCATION: (1)..(12)
;   OTHER INFORMATION: Peptide corresponding to amino acids 367-378 of
;   OTHER INFORMATION: gp38k protein
US-09-459-749D-6

```

```

Query Match          100.0%; Score 21; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 26;
Matches      4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 NLR 4

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 22:05:26 ; Search time 16.8403 Seconds  
 (without alignments)  
 62.382 Million cell updates/sec

Title: US-10-030-735-55  
 Perfect score: 21  
 Sequence: 1 NLRP 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : Published\_Applications\_AA:\*  
 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:\*  
 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:\*  
 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:\*  
 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:\*  
 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result	%	Query				
No.	Score	Match	Length	DB	ID	Description

1	21	100.0	12	9	US-09-459-749D-6	Sequence 6, Appli
2	21	100.0	12	13	US-10-057-113-21	Sequence 21, Appl
3	21	100.0	16	14	US-10-239-313A-590	Sequence 590, App
4	21	100.0	20	14	US-10-115-072-8	Sequence 8, Appli
5	21	100.0	20	14	US-10-115-072-26	Sequence 26, Appl
6	21	100.0	22	10	US-09-974-879-531	Sequence 531, App
7	21	100.0	22	10	US-09-305-736-533	Sequence 533, App
8	21	100.0	22	11	US-09-818-683-533	Sequence 533, App
9	21	100.0	22	12	US-10-621-401-531	Sequence 531, App
10	21	100.0	28	14	US-10-135-919-7	Sequence 7, Appli
11	21	100.0	30	12	US-10-424-599-159585	Sequence 159585,
12	21	100.0	31	15	US-10-264-049-3782	Sequence 3782, Ap
13	21	100.0	36	10	US-09-962-756-1298	Sequence 1298, Ap
14	21	100.0	36	12	US-10-424-599-269387	Sequence 269387,
15	21	100.0	36	15	US-10-253-471-1298	Sequence 1298, Ap
16	21	100.0	36	16	US-10-253-493-1298	Sequence 1298, Ap
17	21	100.0	38	9	US-09-864-761-39711	Sequence 39711, A
18	21	100.0	39	12	US-10-424-599-282214	Sequence 282214,
19	21	100.0	40	9	US-09-864-761-42911	Sequence 42911, A
20	21	100.0	40	12	US-10-424-599-166621	Sequence 166621,
21	21	100.0	40	14	US-10-029-386-28554	Sequence 28554, A
22	21	100.0	41	12	US-10-424-599-252548	Sequence 252548,
23	21	100.0	41	12	US-10-425-114-39972	Sequence 39972, A
24	21	100.0	42	12	US-10-424-599-232796	Sequence 232796,
25	21	100.0	42	12	US-10-424-599-274825	Sequence 274825,
26	21	100.0	44	12	US-10-424-599-172338	Sequence 172338,
27	21	100.0	44	12	US-10-424-599-211767	Sequence 211767,
28	21	100.0	45	12	US-10-282-122A-52283	Sequence 52283, A
29	21	100.0	45	12	US-10-424-599-217108	Sequence 217108,
30	21	100.0	46	12	US-10-424-599-236705	Sequence 236705,
31	21	100.0	46	12	US-10-424-599-280384	Sequence 280384,
32	21	100.0	47	12	US-10-424-599-147128	Sequence 147128,
33	21	100.0	47	15	US-10-264-049-3877	Sequence 3877, Ap
34	21	100.0	48	12	US-10-424-599-199131	Sequence 199131,
35	21	100.0	50	12	US-10-424-599-167146	Sequence 167146,
36	21	100.0	51	14	US-10-106-698-6071	Sequence 6071, Ap
37	21	100.0	52	9	US-09-764-869-721	Sequence 721, App
38	21	100.0	52	12	US-10-424-599-235341	Sequence 235341,
39	21	100.0	52	14	US-10-091-504-721	Sequence 721, App
40	21	100.0	52	15	US-10-227-577-721	Sequence 721, App
41	21	100.0	53	12	US-10-424-599-157355	Sequence 157355,
42	21	100.0	54	12	US-10-424-599-169300	Sequence 169300,
43	21	100.0	54	12	US-10-424-599-180807	Sequence 180807,
44	21	100.0	54	12	US-10-424-599-206555	Sequence 206555,
45	21	100.0	55	12	US-10-424-599-227476	Sequence 227476,
46	21	100.0	56	9	US-09-864-761-44972	Sequence 44972, A
47	21	100.0	56	12	US-10-424-599-207334	Sequence 207334,
48	21	100.0	56	12	US-10-424-599-229568	Sequence 229568,
49	21	100.0	57	10	US-09-764-872-361	Sequence 361, App
50	21	100.0	57	10	US-09-764-891-4940	Sequence 4940, Ap
51	21	100.0	57	14	US-10-205-428-419	Sequence 419, App
52	21	100.0	58	12	US-10-424-599-210299	Sequence 210299,
53	21	100.0	59	12	US-10-424-599-218816	Sequence 218816,
54	21	100.0	59	12	US-10-335-977-7822	Sequence 7822, Ap
55	21	100.0	60	9	US-09-815-242-4871	Sequence 4871, Ap
56	21	100.0	61	9	US-09-895-913A-206	Sequence 206, App

57	21	100.0	61	12	US-10-424-599-156254	Sequence 156254,
58	21	100.0	61	12	US-10-424-599-222932	Sequence 222932,
59	21	100.0	62	12	US-10-282-122A-42399	Sequence 42399, A
60	21	100.0	62	12	US-10-282-122A-57646	Sequence 57646, A
61	21	100.0	62	12	US-10-424-599-150791	Sequence 150791,
62	21	100.0	62	12	US-10-424-599-169636	Sequence 169636,
63	21	100.0	62	12	US-10-424-599-251461	Sequence 251461,
64	21	100.0	63	12	US-10-282-122A-60505	Sequence 60505, A
65	21	100.0	63	12	US-10-424-599-244620	Sequence 244620,
66	21	100.0	64	12	US-10-424-599-271161	Sequence 271161,
67	21	100.0	65	12	US-10-424-599-181478	Sequence 181478,
68	21	100.0	65	12	US-10-424-599-278160	Sequence 278160,
69	21	100.0	65	12	US-10-335-977-7823	Sequence 7823, Ap
70	21	100.0	66	12	US-10-282-122A-46641	Sequence 46641, A
71	21	100.0	66	12	US-10-424-599-157741	Sequence 157741,
72	21	100.0	68	10	US-09-764-891-4443	Sequence 4443, Ap
73	21	100.0	68	12	US-10-424-599-202313	Sequence 202313,
74	21	100.0	69	12	US-10-282-122A-70406	Sequence 70406, A
75	21	100.0	69	12	US-10-282-122A-71233	Sequence 71233, A
76	21	100.0	69	12	US-10-282-122A-72010	Sequence 72010, A
77	21	100.0	69	12	US-10-424-599-163321	Sequence 163321,
78	21	100.0	70	9	US-09-815-242-10523	Sequence 10523, A
79	21	100.0	70	12	US-10-282-122A-52183	Sequence 52183, A
80	21	100.0	70	12	US-10-424-599-208570	Sequence 208570,
81	21	100.0	70	14	US-10-231-417-171	Sequence 171, App
82	21	100.0	72	12	US-10-424-599-144744	Sequence 144744,
83	21	100.0	72	12	US-10-424-599-151322	Sequence 151322,
84	21	100.0	73	12	US-10-424-599-280675	Sequence 280675,
85	21	100.0	74	12	US-10-424-599-155566	Sequence 155566,
86	21	100.0	74	12	US-10-424-599-284175	Sequence 284175,
87	21	100.0	74	14	US-10-106-698-4306	Sequence 4306, Ap
88	21	100.0	74	14	US-10-156-761-12468	Sequence 12468, A
89	21	100.0	75	10	US-09-974-879-529	Sequence 529, App
90	21	100.0	75	10	US-09-305-736-531	Sequence 531, App
91	21	100.0	75	11	US-09-818-683-531	Sequence 531, App
92	21	100.0	75	12	US-10-424-599-180026	Sequence 180026,
93	21	100.0	75	12	US-10-424-599-183205	Sequence 183205,
94	21	100.0	75	12	US-10-424-599-260352	Sequence 260352,
95	21	100.0	75	12	US-10-621-401-529	Sequence 529, App
96	21	100.0	76	9	US-09-925-299-1072	Sequence 1072, Ap
97	21	100.0	76	9	US-09-738-626-4077	Sequence 4077, Ap
98	21	100.0	76	10	US-09-925-299-1072	Sequence 1072, Ap
99	21	100.0	76	12	US-10-282-122A-54169	Sequence 54169, A
100	21	100.0	76	12	US-10-424-599-240808	Sequence 240808,
101	21	100.0	77	9	US-09-815-242-12214	Sequence 12214, A
102	21	100.0	77	9	US-09-815-242-12762	Sequence 12762, A
103	21	100.0	77	12	US-10-282-122A-62267	Sequence 62267, A
104	21	100.0	77	12	US-10-282-122A-62394	Sequence 62394, A
105	21	100.0	77	12	US-10-282-122A-64432	Sequence 64432, A
106	21	100.0	77	12	US-10-424-599-144766	Sequence 144766,
107	21	100.0	77	13	US-10-001-870-172	Sequence 172, App
108	21	100.0	77	14	US-10-113-948-3	Sequence 3, Appli
109	21	100.0	77	14	US-10-106-698-5368	Sequence 5368, Ap
110	21	100.0	78	12	US-10-424-599-161963	Sequence 161963,
111	21	100.0	78	12	US-10-424-599-162677	Sequence 162677,
112	21	100.0	80	12	US-10-424-599-144173	Sequence 144173,
113	21	100.0	80	12	US-10-424-599-157847	Sequence 157847,

114	21	100.0	81	9	US-09-864-761-34627	Sequence 34627, A
115	21	100.0	81	12	US-10-282-122A-63952	Sequence 63952, A
116	21	100.0	82	12	US-10-424-599-250014	Sequence 250014,
117	21	100.0	82	12	US-10-425-114-51140	Sequence 51140, A
118	21	100.0	83	11	US-09-864-408A-5786	Sequence 5786, Ap
119	21	100.0	84	12	US-10-424-599-159444	Sequence 159444,
120	21	100.0	84	12	US-10-424-599-279802	Sequence 279802,
121	21	100.0	85	12	US-10-424-599-193836	Sequence 193836,
122	21	100.0	86	12	US-10-424-599-164070	Sequence 164070,
123	21	100.0	86	12	US-10-424-599-247806	Sequence 247806,
124	21	100.0	86	12	US-10-424-599-276816	Sequence 276816,
125	21	100.0	86	15	US-10-074-978A-343	Sequence 343, App
126	21	100.0	87	12	US-10-424-599-224428	Sequence 224428,
127	21	100.0	87	12	US-10-425-114-45165	Sequence 45165, A
128	21	100.0	89	12	US-10-424-599-282962	Sequence 282962,
129	21	100.0	90	14	US-10-317-832-137	Sequence 137, App
130	21	100.0	91	12	US-10-424-599-231872	Sequence 231872,
131	21	100.0	91	12	US-10-424-599-263273	Sequence 263273,
132	21	100.0	93	12	US-10-424-599-158359	Sequence 158359,
133	21	100.0	93	12	US-10-424-599-266772	Sequence 266772,
134	21	100.0	94	12	US-10-424-599-245073	Sequence 245073,
135	21	100.0	95	12	US-10-424-599-258504	Sequence 258504,
136	21	100.0	95	12	US-10-425-114-72551	Sequence 72551, A
137	21	100.0	96	9	US-09-912-020-264	Sequence 264, App
138	21	100.0	96	14	US-10-284-084-42	Sequence 42, Appl
139	21	100.0	100	12	US-10-424-599-148967	Sequence 148967,
140	21	100.0	101	12	US-10-424-599-243037	Sequence 243037,
141	21	100.0	102	12	US-10-424-599-167287	Sequence 167287,
142	21	100.0	103	14	US-10-106-698-5714	Sequence 5714, Ap
143	21	100.0	105	12	US-10-424-599-165749	Sequence 165749,
144	21	100.0	105	12	US-10-424-599-218454	Sequence 218454,
145	21	100.0	108	9	US-09-738-626-5991	Sequence 5991, Ap
146	21	100.0	109	12	US-10-282-122A-55681	Sequence 55681, A
147	21	100.0	109	12	US-10-424-599-256986	Sequence 256986,
148	21	100.0	109	14	US-10-219-220-144	Sequence 144, App
149	21	100.0	110	12	US-10-424-599-241032	Sequence 241032,
150	21	100.0	111	12	US-10-424-599-199937	Sequence 199937,

#### ALIGNMENTS

#### RESULT 1

US-09-459-749D-6

; Sequence 6, Application US/09459749D

; Patent No. US20020136716A1

; GENERAL INFORMATION:

; APPLICANT: Millis, Albert J. T.

; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration

; FILE REFERENCE: 0794.016A

; CURRENT APPLICATION NUMBER: US/09/459,749D

; CURRENT FILING DATE: 1999-12-10

; PRIOR APPLICATION NUMBER: 60/111,856

; PRIOR FILING DATE: 1998-12-11

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6



```

;   LENGTH: 12
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:SITE
;   NAME/KEY: SITE
;   LOCATION: (1)..(12)
;   OTHER INFORMATION: Peptide corresponding to amino acids 367-378 of
;   OTHER INFORMATION: gp38k protein
US-09-459-749D-6

```

```

Query Match          100.0%;   Score 21;   DB 9;   Length 12;
Best Local Similarity 100.0%;   Pred. No. 70;
Matches      4;   Conservative    0;   Mismatches    0;   Indels    0;   Gaps    0;

```

```

Qy      1 NLRf 4
        ||||
Db      1 NLRf 4

```

# RESULT 2

US-10-057-113-21

; Sequence 21, Application US/10057113

; Publication No. US20020110909A1

## ; GENERAL INFORMATION:

```

;   APPLICANT: THE UNITED STATES OF AMERICA, AS REPRESENTED
;               BY THE SECRETARY, DEPARTMENT OF HEALTH AND HUMAN SERVICES IN
;               CARE OF NATIONAL INSTITUTES OF HEALTH Office of Technology
;               Transfer Suite 325, 6011 Executive Boulevard, Rockville, MD
;               20852

```

; TITLE OF INVENTION: GLYCOPROTEIN HORMONE SUPERAGONISTS

; NUMBER OF SEQUENCES: 37

## ; CORRESPONDENCE ADDRESS:

```

;   ADDRESSEE: NEEDLE & ROSENBERG, P.C.
;   STREET: Suite 1200, 127 Peachtree Street
;   CITY: Atlanta
;   STATE: Georgia
;   COUNTRY: USA
;   ZIP: 30303

```

## ; COMPUTER READABLE FORM:

```

;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30

```

## ; CURRENT APPLICATION DATA:

```

;   APPLICATION NUMBER: US/10/057,113
;   FILING DATE: 25-Jan-2002
;   CLASSIFICATION: <Unknown>

```

## ; PRIOR APPLICATION DATA:

```

;   APPLICATION NUMBER: PCT/US96/06483
;   FILING DATE: <Unknown>

```

## ; ATTORNEY/AGENT INFORMATION:

```

;   NAME: Perryman, David G.
;   REGISTRATION NUMBER: 33,438
;   REFERENCE/DOCKET NUMBER: 14014.0185/P

```

## ; TELECOMMUNICATION INFORMATION:

```

;   TELEPHONE: 404/688-0770

```

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:11:45 ; Search time 5.10924 Seconds  
 (without alignments)  
 75.308 Million cell updates/sec

Title: US-10-030-735-55  
 Perfect score: 21  
 Sequence: 1 NLRP 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		Length	DB	ID	Description
		Match					
1	21	100.0		28	2	B60583	glycoprotein hormo
2	21	100.0		43	2	T44391	ribosomal protein
3	21	100.0		50	2	S11603	hypothetical prote
4	21	100.0		62	2	T18774	hypothetical prote
5	21	100.0		63	2	A87405	ribosomal protein
6	21	100.0		63	2	AG1778	ribosomal protein
7	21	100.0		63	2	AH1402	ribosomal protein
8	21	100.0		64	2	H90959	hypothetical prote
9	21	100.0		66	1	R5BS29	ribosomal protein
10	21	100.0		66	1	R5BS2L	ribosomal protein
11	21	100.0		66	2	AG3347	LSU ribosomal prot
12	21	100.0		66	2	AH2814	50S ribosomal prot
13	21	100.0		66	2	C97593	50S ribosomal prot

14	21	100.0	67	2	F83667	ribosomal protein
15	21	100.0	69	2	E90021	50S ribosomal prot
16	21	100.0	69	2	A71363	hypothetical prote
17	21	100.0	71	2	A71671	ribosomal protein
18	21	100.0	71	2	F97824	50S ribosomal prot
19	21	100.0	77	2	H70642	probable ribosomal
20	21	100.0	77	2	T17858	hypothetical prote
21	21	100.0	79	2	B90806	hypothetical prote
22	21	100.0	79	2	C69036	hypothetical prote
23	21	100.0	80	2	T45372	ribosomal protein
24	21	100.0	81	2	A87141	50S ribosomal prot
25	21	100.0	82	2	S71166	RNA-directed DNA p
26	21	100.0	82	2	E86562	natural UGA frame-
27	21	100.0	82	2	T15818	hypothetical prote
28	21	100.0	86	2	T07204	hypothetical prote
29	21	100.0	87	2	S38681	major histocompati
30	21	100.0	87	2	S38682	major histocompati
31	21	100.0	87	2	S38685	major histocompati
32	21	100.0	87	2	S03117	class II histocomp
33	21	100.0	88	2	D47759	retrovirus-related
34	21	100.0	89	2	S38676	MHC class II histo
35	21	100.0	95	2	G71630	integration host f
36	21	100.0	96	2	E90894	hypothetical prote
37	21	100.0	96	2	D85723	hypothetical prote
38	21	100.0	96	2	AF0050	conserved hypothet
39	21	100.0	96	2	A64906	hypothetical prote
40	21	100.0	97	2	S20287	lutropin alpha cha
41	21	100.0	97	2	G87667	hypothetical prote
42	21	100.0	102	2	AB0101	hypothetical prote
43	21	100.0	102	2	G72700	hypothetical prote
44	21	100.0	104	2	D95916	hypothetical prote
45	21	100.0	105	2	T38761	hypothetical prote
46	21	100.0	108	2	S66828	hypothetical prote
47	21	100.0	109	2	AF0940	conserved hypothet
48	21	100.0	112	2	S64730	protein secretion
49	21	100.0	112	2	C71221	hypothetical prote
50	21	100.0	112	2	AF2549	hypothetical prote
51	21	100.0	120	2	G86654	transcription regu
52	21	100.0	121	2	A46561	MHC class II histo
53	21	100.0	129	2	AI0204	chemotaxis protein
54	21	100.0	129	2	H97938	hypothetical prote
55	21	100.0	130	2	T41117	probable Cytochrom
56	21	100.0	132	2	T50108	yeast APG12 Protei
57	21	100.0	133	2	E69338	hypothetical prote
58	21	100.0	137	2	A87586	hypothetical prote
59	21	100.0	140	2	AB2692	transcription regu
60	21	100.0	140	2	F97473	heavy metal depend
61	21	100.0	142	2	T25514	hypothetical prote
62	21	100.0	144	2	E90120	60S ribosomal prot
63	21	100.0	147	2	B81415	50S ribosomal prot
64	21	100.0	149	2	AE3095	conserved hypothet
65	21	100.0	149	2	E98191	hypothetical prote
66	21	100.0	150	2	F69292	nigerythrin homolo
67	21	100.0	153	2	S76676	hypothetical prote
68	21	100.0	156	2	AH1363	hypothetical prote
69	21	100.0	161	2	A69031	conserved hypothet
70	21	100.0	165	2	F70161	ribosomal protein

71	21	100.0	167	1	S48484	signal peptidase (
72	21	100.0	172	2	S28683	hypothetical prote
73	21	100.0	173	2	T21710	hypothetical prote
74	21	100.0	179	2	T36855	conserved hypothet
75	21	100.0	183	2	A69008	tetrahydromethanop
76	21	100.0	183	2	B84785	hypothetical prote
77	21	100.0	184	2	A81561	hypothetical prote
78	21	100.0	184	2	C72110	hypothetical prote
79	21	100.0	184	2	D86513	hypothetical prote
80	21	100.0	184	2	G87480	hypothetical prote
81	21	100.0	185	2	I39736	hypothetical prote
82	21	100.0	185	2	AE1902	hypothetical prote
83	21	100.0	186	2	G86154	hypothetical prote
84	21	100.0	187	2	T06581	probable deoxyribo
85	21	100.0	187	2	B82857	pre-pilin like lea
86	21	100.0	188	2	D97114	n6-adenine-specifi
87	21	100.0	192	2	A45836	MHC class II histo
88	21	100.0	193	2	A36672	transcription init
89	21	100.0	196	2	AD2304	hypothetical prote
90	21	100.0	197	2	T09257	late embryonic abu
91	21	100.0	201	2	S69748	hypothetical prote
92	21	100.0	203	2	F75554	MutT/nudix family
93	21	100.0	205	2	T20128	hypothetical prote
94	21	100.0	205	2	E70199	competence protein
95	21	100.0	207	2	C82125	heme exporter prot
96	21	100.0	208	2	T05886	hypothetical prote
97	21	100.0	208	2	AC0420	PhnH protein [limpo
98	21	100.0	209	2	C87530	uracil phosphoribo
99	21	100.0	212	2	T51646	probable transcrip
100	21	100.0	212	2	T06015	hypothetical prote
101	21	100.0	213	2	AE2906	hypothetical prote
102	21	100.0	213	2	F97681	ABC-type transport
103	21	100.0	213	2	C69641	ATP phosphoribosyl
104	21	100.0	213	2	T47382	hypothetical prote
105	21	100.0	216	2	S15223	ctrD protein - Nei
106	21	100.0	216	2	F82013	capsule polysaccha
107	21	100.0	216	2	H81241	capsule polysaccha
108	21	100.0	217	1	BVHIXA	capsulation protei
109	21	100.0	219	2	F97289	probable membrane
110	21	100.0	219	2	E70431	hypothetical prote
111	21	100.0	225	2	T52011	ethylene responsiv
112	21	100.0	226	2	G69129	hypothetical prote
113	21	100.0	227	2	G89807	exotoxin 14 [impor
114	21	100.0	229	2	B90534	trna/rrna methyltr
115	21	100.0	229	2	AG2390	hypothetical prote
116	21	100.0	232	1	HLMSE2	H-2 class II histo
117	21	100.0	233	2	I55654	MHC class II prote
118	21	100.0	234	2	PC4398	glutamine-tRNA lig
119	21	100.0	237	2	A21200	H-2 class II histo
120	21	100.0	237	2	T46121	hypothetical prote
121	21	100.0	241	2	I48657	I-E(b-beta) protei
122	21	100.0	245	2	B64007	hypothetical prote
123	21	100.0	246	2	T10127	transcription init
124	21	100.0	247	2	A81388	3-oxoacyl-[acyl-ca
125	21	100.0	248	2	D90986	hypothetical prote
126	21	100.0	248	2	G85831	hypothetical prote
127	21	100.0	248	2	F64971	hypothetical prote

128	21	100.0	248	2	AD0769	probable glycosylt
129	21	100.0	249	2	E75193	hypothetical prote
130	21	100.0	249	2	AF3225	hypothetical prote
131	21	100.0	250	2	F81961	probable hydroxyac
132	21	100.0	250	2	B86232	hypothetical prote
133	21	100.0	252	2	A95071	amino acid ABC tra
134	21	100.0	252	2	H81017	hydroxyacylglutath
135	21	100.0	252	2	AB1748	acetylerase hom
136	21	100.0	252	2	AI1378	acetylerase hom
137	21	100.0	255	2	S04668	hypothetical prote
138	21	100.0	256	2	AI3416	hypothetical membr
139	21	100.0	257	1	S40985	H+-transporting tw
140	21	100.0	258	2	D75520	conserved hypothet
141	21	100.0	258	2	G83417	probable hydroxyac
142	21	100.0	258	2	S77158	hydrogenase access
143	21	100.0	260	2	T15814	hypothetical prote
144	21	100.0	261	2	D70112	glutamate racemase
145	21	100.0	261	2	A69142	hypothetical prote
146	21	100.0	261	2	C84584	probable cAMP-depe
147	21	100.0	262	2	H45734	3'-aminoglycoside
148	21	100.0	262	2	C86776	kanamycin kinase (
149	21	100.0	262	2	T06600	acetyl-CoA carboxy
150	21	100.0	263	2	T16329	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

B60583

glycoprotein hormones alpha chain - bullfrog (fragment)

N;Alternate names: follitropin alpha chain; gonadotropin alpha chain; lutropin alpha chain; thyrotropin alpha chain

C;Species: Rana catesbeiana (bullfrog)

C;Date: 17-Apr-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999

C;Accession: B60583

R;Bergert, E.R.; Madden, B.; McCormick, D.J.; Papkoff, H.; Ryan, R.J.

Endocrinology 127, 2985-2989, 1990

A;Title: The antigenic structure of the human glycoprotein hormone alpha-subunit: Cross-species comparisons.

A;Reference number: A60583; MUID:91065269; PMID:1701134

A;Accession: B60583

A;Molecule type: protein

A;Residues: 1-28 <BER>

C;Superfamily: glycoprotein hormones alpha chain

C;Keywords: glycoprotein; hormone

Query Match 100.0%; Score 21; DB 2; Length 28;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4

||||

Db 19 NLRf 22

##### RESULT 2

T44391  
 ribosomal protein L16 [imported] - *Bacillus halodurans*  
 C;Species: *Bacillus halodurans*  
 C;Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
 C;Accession: T44391  
 R;Takami, H.; Takaki, Y.; Nakasone, K.; Hiramata, C.; Inoue, A.; Horikoshi, K.  
 Biosci. Biotechnol. Biochem. 63, 452-455, 1999  
 A;Title: Sequence analysis of a 32-kb region including the major ribosomal  
 protein gene clusters from alkaliphilic *Bacillus* sp. strain C-125.  
 A;Reference number: Z22756; MUID:99209008; PMID:10192928  
 A;Accession: T44391  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-43 <TAK>  
 A;Cross-references: EMBL:AB017508; NID:g4512395; PIDN:BAA75279.1; PID:g4512412  
 A;Experimental source: strain C-125  
 C;Genetics:  
 A;Gene: rpmC  
 C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
 ||||  
 Db 3 NLRF 6

#### RESULT 3

S11603  
 hypothetical protein 2 (16S rRNA 3' region) - *Thermoplasma acidophilum*  
 (fragment)  
 C;Species: *Thermoplasma acidophilum*  
 C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1993  
 C;Accession: S11603  
 R;Ree, H.K.; Cao, K.; Thurlow, D.L.; Zimmermann, R.A.  
 Can. J. Microbiol. 35, 124-133, 1989  
 A;Title: The structure and organization of the 16S ribosomal RNA gene from the  
 archaeobacterium *Thermoplasma acidophilum*.  
 A;Reference number: S11602; MUID:89248669; PMID:2470478  
 A;Accession: S11603  
 A;Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 1-50 <REE>

Query Match 100.0%; Score 21; DB 2; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
 ||||  
 Db 27 NLRF 30

#### RESULT 4

T18774

OM protein - protein search, using sw model

Run on: April 7, 2004, 18:58:36 ; Search time 2.72269 Seconds  
 (without alignments)  
 76.498 Million cell updates/sec

Title: US-10-030-735-53  
 Perfect score: 20  
 Sequence: 1 QVRF 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	20	100.0	15	1 ALLS_MANSE	P42559 manduca sex
2	20	100.0	63	1 YM26_MARPO	P38468 marchantia
3	20	100.0	81	1 Y65_BPT3	P20329 bacteriopha
4	20	100.0	84	1 Y65_BPT7	P03800 bacteriopha
5	20	100.0	102	1 Y66A_HAEIN	O86228 haemophilus
6	20	100.0	114	1 RS13_MUSDO	P27072 musca domes
7	20	100.0	119	1 SMPA_VIBCH	P52117 vibrio chol
8	20	100.0	134	1 VAL3_CLVK	P14977 cassava lat
9	20	100.0	134	1 VAL3_CLVN	P14969 cassava lat
10	20	100.0	140	1 COX6_SCHPO	Q9utf6 schizosacch
11	20	100.0	144	1 PSB1_ECOLI	P10031 escherichia
12	20	100.0	144	1 PSB2_ECOLI	P10032 escherichia
13	20	100.0	150	1 RS13_ANOGA	P52811 anopheles g
14	20	100.0	150	1 RS13_DROME	Q03334 drosophila
15	20	100.0	150	1 RS13_HUMAN	Q02546 homo sapien
16	20	100.0	150	1 RS13 ICTPU	P47772 ictalurus p
17	20	100.0	150	1 RS13_SCHPO	P28189 schizosacch

18	20	100.0	150	1	RS13_XENLA	P49393	xenopus lae
19	20	100.0	151	1	RS13_AGABI	P78571	agaricus bi
20	20	100.0	153	1	RECX_VIBVU	Q8dc50	vibrio vuln
21	20	100.0	155	1	RECX_VIBPA	Q87lr2	vibrio para
22	20	100.0	159	1	RECX_RALSO	Q8yly5	ralstonia s
23	20	100.0	173	1	INSJ_ECOLI	P19768	escherichia
24	20	100.0	173	1	INSJ_SHIFL	P60046	shigella fl
25	20	100.0	175	1	OMLA_PSEFL	O68564	pseudomonas
26	20	100.0	176	1	OMLA_PSEAE	O68562	pseudomonas
27	20	100.0	184	1	BGB_DROME	Q24040	drosophila
28	20	100.0	190	1	GCH1_CAMJE	P51594	campylobact
29	20	100.0	197	1	GCH1_RHILO	Q981q6	rhizobium l
30	20	100.0	204	1	GCH1_RHIME	Q92qb4	rhizobium m
31	20	100.0	205	1	ALKH_ZYMMO	Q00384	z khg/kdpg
32	20	100.0	207	1	YL96_RALSO	P58634	ralstonia s
33	20	100.0	208	1	GCH1_AGRT5	Q8uek8	agrobacteri
34	20	100.0	212	1	HIS5_PSEPK	Q88r44	pseudomonas
35	20	100.0	213	1	ALKH_ECOLI	P10177	e khg/kdpg
36	20	100.0	213	1	GCH1_BRUME	Q8yh94	brucella me
37	20	100.0	213	1	GCH1_BRUSU	Q8g0l4	brucella su
38	20	100.0	213	1	HI51_PSEAE	Q9hu42	pseudomonas
39	20	100.0	215	1	YAA4_YEAST	P39703	saccharomyc
40	20	100.0	224	1	YHCA_ECOLI	P28722	escherichia
41	20	100.0	229	1	UCK_CAEEL	Q17413	caenorhabdi
42	20	100.0	231	1	RL1_PSEAE	Q9hwc6	pseudomonas
43	20	100.0	239	1	LPXH_PASMU	Q9cpe3	pasteurella
44	20	100.0	243	1	YGGJ_ECOLI	P37912	escherichia
45	20	100.0	246	1	VATD_MANSE	Q9u0s4	manduca sex
46	20	100.0	266	1	YAFD_ECOLI	P30865	escherichia
47	20	100.0	266	1	YAFD_SALTI	Q8z985	salmonella
48	20	100.0	266	1	YAFD_SALTY	Q8zrm4	salmonella
49	20	100.0	267	1	PPNK_LACPL	Q88v61	lactobacill
50	20	100.0	269	1	MURI_PASMU	P57887	pasteurella
51	20	100.0	269	1	PYR5_FREDI	P11401	fremyella d
52	20	100.0	270	1	HMPH_HUMAN	Q03014	homo sapien
53	20	100.0	271	1	HMPH_MOUSE	P43120	mus musculu
54	20	100.0	271	1	PYR6_FREDI	P14880	fremyella d
55	20	100.0	275	1	BLOI_PSEAE	O07293	pseudomonas
56	20	100.0	277	1	HMPH_CHICK	Q05502	gallus gall
57	20	100.0	278	1	PYR2_MASLA	P11399	mastigoclad
58	20	100.0	285	1	YD98_YEAST	Q03835	saccharomyc
59	20	100.0	296	1	RL5_MOUSE	P47962	mus musculu
60	20	100.0	296	1	RL5_RAT	P09895	rattus norv
61	20	100.0	325	1	Y272_SYNY3	P73893	synechocyst
62	20	100.0	327	1	RL5_ANOGA	O44248	anopheles g
63	20	100.0	327	1	SYFA_PASMU	P57860	pasteurella
64	20	100.0	328	1	Y002_NPVAC	P24655	autographa
65	20	100.0	329	1	SRA8_CAEEL	Q09210	caenorhabdi
66	20	100.0	329	1	SYFA_HAEIN	P43819	haemophilus
67	20	100.0	330	1	SYFA_NEIMA	Q9jrr76	neisseria m
68	20	100.0	330	1	SYFA_NEIMB	Q9k092	neisseria m
69	20	100.0	333	1	RRS1_CAEEL	Q9xvt0	caenorhabdi
70	20	100.0	336	1	YERI_BACSU	O34640	bacillus su
71	20	100.0	339	1	TCMO_STRGA	P39896	streptomyce
72	20	100.0	340	1	MOAA_STACA	Q9zim6	staphylococ
73	20	100.0	340	1	REP4_ECOLI	P13961	escherichia
74	20	100.0	343	1	REP7_ECOLI	Q52221	escherichia



75	20	100.0	343	1	REP8_ECOLI	Q52346	escherichia
76	20	100.0	343	1	REPZ_ECOLI	P18023	escherichia
77	20	100.0	344	1	SYFA_RALSO	Q8xz25	ralstonia s
78	20	100.0	353	1	MURG_HELPJ	Q9zk59	helicobacte
79	20	100.0	353	1	MURG_HELPY	O25770	helicobacte
80	20	100.0	353	1	PMA1_MOUSE	Q8c1c8	mus musculu
81	20	100.0	353	1	PMA1_RAT	Q8vhz4	rattus norv
82	20	100.0	365	1	FLGI_ECO57	P58203	escherichia
83	20	100.0	365	1	FLGI_ECOLI	P75941	escherichia
84	20	100.0	365	1	FLGI_SALTY	P15930	salmonella
85	20	100.0	367	1	DP3B_PSEAE	Q9i7c4	pseudomonas
86	20	100.0	367	1	FLGI_SALTI	Q8z7k2	salmonella
87	20	100.0	367	1	MTC1_CHVI3	P36216	chlorella v
88	20	100.0	369	1	FLGI_YERPE	Q8zfb1	yersinia pe
89	20	100.0	373	1	NSDL_HUMAN	Q15738	homo sapien
90	20	100.0	378	1	TRMU_XANAC	Q8pl08	xanthomonas
91	20	100.0	381	1	SSUD_PSEAE	Q9hyg2	pseudomonas
92	20	100.0	382	1	CRTY_PANAN	P21687	pantoea ana
93	20	100.0	390	1	ASSY_SULTO	Q970v0	sulfolobus
94	20	100.0	395	1	ASSY_RHOFA	Q93jq8	rhodococcus
95	20	100.0	396	1	ASSY_STRA3	Q8e7n1	streptococc
96	20	100.0	396	1	ASSY_STRA5	Q8e272	streptococc
97	20	100.0	396	1	ASSY_STRMU	Q8cwz0	streptococc
98	20	100.0	396	1	RF1M_KLULA	P41767	kluyveromyc
99	20	100.0	397	1	ASSY_PYRAE	Q8zu97	pyrobaculum
100	20	100.0	397	1	ASSY_STRCL	P50986	streptomyce
101	20	100.0	397	1	ASSY_THEVO	Q97a55	thermoplasm
102	20	100.0	397	1	ASSY_XANCP	Q8p8j4	xanthomonas
103	20	100.0	398	1	ASSY_CLOTE	P59602	clostridium
104	20	100.0	398	1	ASSY_LACLA	P57799	lactococcus
105	20	100.0	398	1	ASSY_MYCTU	P94993	mycobacteri
106	20	100.0	398	1	ASSY_STRR6	Q8dri5	streptococc
107	20	100.0	399	1	ASSY_MYCLE	Q9cc10	mycobacteri
108	20	100.0	400	1	ASSY_ANASP	Q8ymx6	anabaena sp
109	20	100.0	400	1	ASSY_CLOAB	Q97ke6	clostridium
110	20	100.0	400	1	ASSY_SYNY3	P77973	synechocyst
111	20	100.0	400	1	ASSY_THETH	P59846	thermus the
112	20	100.0	401	1	ASSY_AQUAE	O67213	aquifex aeo
113	20	100.0	401	1	ASSY_BACAA	Q81kv7	bacillus an
114	20	100.0	401	1	ASSY_BACCR	Q817c6	bacillus ce
115	20	100.0	401	1	ASSY_BUCAP	Q8ka60	buchnera ap
116	20	100.0	401	1	ASSY_CHLTE	Q8kde0	chlorobium
117	20	100.0	401	1	ASSY_COREF	Q8ftm9	corynebacte
118	20	100.0	401	1	ASSY_CORGL	O85176	corynebacte
119	20	100.0	401	1	ASSY_STAAM	Q99vc7	staphylococ
120	20	100.0	401	1	ASSY_STAAW	Q8nxf2	staphylococ
121	20	100.0	401	1	ASSY_STAEP	Q8cpu3	staphylococ
122	20	100.0	401	1	ASSY_SYNEL	Q8dky7	synechococc
123	20	100.0	401	1	ASSY_XYLFA	Q9pem9	xylella fas
124	20	100.0	401	1	ASSY_XYLFT	P59606	xylella fas
125	20	100.0	402	1	ASSY_DEIRA	Q9rwj4	deinococcus
126	20	100.0	403	1	ASSY_BACSU	O34347	bacillus su
127	20	100.0	403	1	ASSY_BUCAI	P57158	buchnera ap
128	20	100.0	403	1	ASSY_LEPIN	Q8eyp7	leptospira
129	20	100.0	404	1	ASSY_LISIN	Q929s9	listeria in
130	20	100.0	404	1	ASSY_LISMO	Q8y5h2	listeria mo
131	20	100.0	404	1	ASSY_NITEU	Q82up5	nitrosomona

132	20	100.0	404	1	ASSY_STRAW	Q827z1	streptomyce
133	20	100.0	404	1	ASSY_VIBCH	Q9knt8	vibrio chol
134	20	100.0	404	1	ASSY_VIBPA	P59605	vibrio para
135	20	100.0	405	1	ASSY_CLOPE	Q8xmj7	clostridium
136	20	100.0	405	1	ASSY_PSEAE	Q9hy84	pseudomonas
137	20	100.0	405	1	ASSY_PSEPK	P59604	pseudomonas
138	20	100.0	405	1	ASSY_PSESM	Q87xm3	pseudomonas
139	20	100.0	405	1	ASSY_RHIME	Q92l73	rhizobium m
140	20	100.0	405	1	ASSY_THEAC	Q9hkf1	thermoplasm
141	20	100.0	405	1	ASSY_VIBVU	Q8dcn0	vibrio vuln
142	20	100.0	406	1	ASSY_BRUME	Q8yek8	brucella me
143	20	100.0	406	1	ASSY_BRUSU	Q8g376	brucella su
144	20	100.0	406	1	ASSY_CAMJE	Q9phk7	campylobact
145	20	100.0	406	1	ASSY_SULSO	Q9ux31	sulfolobus
146	20	100.0	407	1	ASSY_AGRT5	Q8uc31	agrobacteri
147	20	100.0	407	1	ASSY_RHILO	Q98e81	rhizobium l
148	20	100.0	407	1	ASSY_SHEON	Q8ek28	shewanella
149	20	100.0	408	1	ASSY_CAUCR	Q9abul	caulobacter
150	20	100.0	408	1	ASSY_HELMO	Q8gdu2	heliobacill

#### ALIGNMENTS

##### RESULT 1

##### ALLS\_MANSE

ID ALLS\_MANSE STANDARD; PRT; 15 AA.  
AC P42559;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Allatostatin (Mas-AS).  
OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Sphingioidea;  
OC Sphingidae; Sphinginae; Manduca.  
OX NCBI\_TaxID=7130;  
RN [1]  
RP SEQUENCE.  
RC TISSUE=Head;  
RX MEDLINE=92052112; PubMed=1946359;  
RA Kramer S.J., Toschi A., Miller C.A., Kataoka H., Quistad G.B.,  
RA Li J.P., Carney R.L., Schooley D.A.;  
RT "Identification of an allatostatin from the tobacco hornworm Manduca  
RT sexta.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:9458-9462(1991).  
CC -!- FUNCTION: Strongly inhibits juvenile hormone biosynthesis in vitro  
CC by the corpora allata from fifth-stadium larvae and adult females.  
CC -!- SIMILARITY: Belongs to the allatostatin family.  
DR PIR; A61612; A61612.  
KW Neuropeptide; Pyrrolidone carboxylic acid.  
FT MOD\_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.  
SQ SEQUENCE 15 AA; 1908 MW; 1605B77CDEBC838E CRC64;

Query Match 100.0%; Score 20; DB 1; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.6;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVRF 4  
    ||||  
Db 1 QVRF 4

RESULT 2

YM26\_MARPO

ID YM26\_MARPO STANDARD; PRT; 63 AA.  
AC P38468;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 7.1 kDa protein in NAD2 3'region (ORF 63).  
GN YMF26.  
OS Marchantia polymorpha (Liverwort).  
OG Mitochondrion.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
OC Marchantiopsida; Marchantiidae; Marchantiales; Marchantiineae;  
OC Marchantiaceae; Marchantia.  
OX NCBI\_TaxID=3197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92114051; PubMed=1731062;  
RA Oda K., Yamato K., Ohta E., Nakamura Y., Takemura M., Nozato N.,  
RA Akashi K., Kanegae T., Ogura Y., Kohchi T., Ohyama K.;  
RT "Gene organization deduced from the complete sequence of liverwort  
RT Marchantia polymorpha mitochondrial DNA. A primitive form of plant  
RT mitochondrial genome.";  
RL J. Mol. Biol. 223:1-7(1992).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M68929; AAC09400.1; -.  
DR PIR; S25962; S25962.  
KW Mitochondrion; Hypothetical protein.  
SQ SEQUENCE 63 AA; 7143 MW; 607FCCB01A9FDA2A CRC64;

Query Match 100.0%; Score 20; DB 1; Length 63;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QVRF 4  
    ||||  
Db 6 QVRF 9

RESULT 3

Y65\_BPT3

ID Y65\_BPT3 STANDARD; PRT; 81 AA.  
AC P20329;

OM protein - protein search, using sw model

Run on: April 7, 2004, 19:11:45 ; Search time 5.10924 Seconds  
 (without alignments)  
 75.308 Million cell updates/sec

Title: US-10-030-735-55  
 Perfect score: 21  
 Sequence: 1 NLRP 4

Scoring table: BLOSUM62  
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
 Maximum Match 100%  
 Listing first 150 summaries

Database : PIR\_78:\*  
 1: pir1:\*  
 2: pir2:\*  
 3: pir3:\*  
 4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	% Query		DB	ID	Description
		Match	Length			
1	21	100.0	28	2	B60583	glycoprotein hormo
2	21	100.0	43	2	T44391	ribosomal protein
3	21	100.0	50	2	S11603	hypothetical prote
4	21	100.0	62	2	T18774	hypothetical prote
5	21	100.0	63	2	A87405	ribosomal protein
6	21	100.0	63	2	AG1778	ribosomal protein
7	21	100.0	63	2	AH1402	ribosomal protein
8	21	100.0	64	2	H90959	hypothetical prote
9	21	100.0	66	1	R5BS29	ribosomal protein
10	21	100.0	66	1	R5BS2L	ribosomal protein
11	21	100.0	66	2	AG3347	LSU ribosomal prot
12	21	100.0	66	2	AH2814	50S ribosomal prot
13	21	100.0	66	2	C97593	50S ribosomal prot

14	21	100.0	67	2	F83667	ribosomal protein
15	21	100.0	69	2	E90021	50S ribosomal prot
16	21	100.0	69	2	A71363	hypothetical prote
17	21	100.0	71	2	A71671	ribosomal protein
18	21	100.0	71	2	F97824	50S ribosomal prot
19	21	100.0	77	2	H70642	probable ribosomal
20	21	100.0	77	2	T17858	hypothetical prote
21	21	100.0	79	2	B90806	hypothetical prote
22	21	100.0	79	2	C69036	hypothetical prote
23	21	100.0	80	2	T45372	ribosomal protein
24	21	100.0	81	2	A87141	50S ribosomal prot
25	21	100.0	82	2	S71166	RNA-directed DNA p
26	21	100.0	82	2	E86562	natural UGA frame-
27	21	100.0	82	2	T15818	hypothetical prote
28	21	100.0	86	2	T07204	hypothetical prote
29	21	100.0	87	2	S38681	major histocompati
30	21	100.0	87	2	S38682	major histocompati
31	21	100.0	87	2	S38685	major histocompati
32	21	100.0	87	2	S03117	class II histocomp
33	21	100.0	88	2	D47759	retrovirus-related
34	21	100.0	89	2	S38676	MHC class II histo
35	21	100.0	95	2	G71630	integration host f
36	21	100.0	96	2	E90894	hypothetical prote
37	21	100.0	96	2	D85723	hypothetical prote
38	21	100.0	96	2	AF0050	conserved hypothet
39	21	100.0	96	2	A64906	hypothetical prote
40	21	100.0	97	2	S20287	lutropin alpha cha
41	21	100.0	97	2	G87667	hypothetical prote
42	21	100.0	102	2	AB0101	hypothetical prote
43	21	100.0	102	2	G72700	hypothetical prote
44	21	100.0	104	2	D95916	hypothetical prote
45	21	100.0	105	2	T38761	hypothetical prote
46	21	100.0	108	2	S66828	hypothetical prote
47	21	100.0	109	2	AF0940	conserved hypothet
48	21	100.0	112	2	S64730	protein secretion
49	21	100.0	112	2	C71221	hypothetical prote
50	21	100.0	112	2	AF2549	hypothetical prote
51	21	100.0	120	2	G86654	transcription regu
52	21	100.0	121	2	A46561	MHC class II histo
53	21	100.0	129	2	AI0204	chemotaxis protein
54	21	100.0	129	2	H97938	hypothetical prote
55	21	100.0	130	2	T41117	probable Cytochrom
56	21	100.0	132	2	T50108	yeast APG12 Protei
57	21	100.0	133	2	E69338	hypothetical prote
58	21	100.0	137	2	A87586	hypothetical prote
59	21	100.0	140	2	AB2692	transcription regu
60	21	100.0	140	2	F97473	heavy metal depend
61	21	100.0	142	2	T25514	hypothetical prote
62	21	100.0	144	2	E90120	60S ribosomal prot
63	21	100.0	147	2	B81415	50S ribosomal prot
64	21	100.0	149	2	AE3095	conserved hypothet
65	21	100.0	149	2	E98191	hypothetical prote
66	21	100.0	150	2	F69292	nigerythrin homolo
67	21	100.0	153	2	S76676	hypothetical prote
68	21	100.0	156	2	AH1363	hypothetical prote
69	21	100.0	161	2	A69031	conserved hypothet
70	21	100.0	165	2	F70161	ribosomal protein

71	21	100.0	167	1	S48484	signal peptidase (
72	21	100.0	172	2	S28683	hypothetical prote
73	21	100.0	173	2	T21710	hypothetical prote
74	21	100.0	179	2	T36855	conserved hypothet
75	21	100.0	183	2	A69008	tetrahydromethanop
76	21	100.0	183	2	B84785	hypothetical prote
77	21	100.0	184	2	A81561	hypothetical prote
78	21	100.0	184	2	C72110	hypothetical prote
79	21	100.0	184	2	D86513	hypothetical prote
80	21	100.0	184	2	G87480	hypothetical prote
81	21	100.0	185	2	I39736	hypothetical prote
82	21	100.0	185	2	AE1902	hypothetical prote
83	21	100.0	186	2	G86154	hypothetical prote
84	21	100.0	187	2	T06581	probable deoxyribo
85	21	100.0	187	2	B82857	pre-pilin like lea
86	21	100.0	188	2	D97114	n6-adenine-specifi
87	21	100.0	192	2	A45836	MHC class II histo
88	21	100.0	193	2	A36672	transcription init
89	21	100.0	196	2	AD2304	hypothetical prote
90	21	100.0	197	2	T09257	late embryonic abu
91	21	100.0	201	2	S69748	hypothetical prote
92	21	100.0	203	2	F75554	MutT/nudix family
93	21	100.0	205	2	T20128	hypothetical prote
94	21	100.0	205	2	E70199	competence protein
95	21	100.0	207	2	C82125	heme exporter prot
96	21	100.0	208	2	T05886	hypothetical prote
97	21	100.0	208	2	AC0420	PhnH protein limpo
98	21	100.0	209	2	C87530	uracil phosphoribo
99	21	100.0	212	2	T51646	probable transcrip
100	21	100.0	212	2	T06015	hypothetical prote
101	21	100.0	213	2	AE2906	hypothetical prote
102	21	100.0	213	2	F97681	ABC-type transport
103	21	100.0	213	2	C69641	ATP phosphoribosyl
104	21	100.0	213	2	T47382	hypothetical prote
105	21	100.0	216	2	S15223	ctrD protein - Nei
106	21	100.0	216	2	F82013	capsule polysaccha
107	21	100.0	216	2	H81241	capsule polysaccha
108	21	100.0	217	1	BVHIXA	capsulation protei
109	21	100.0	219	2	F97289	probable membrane
110	21	100.0	219	2	E70431	hypothetical prote
111	21	100.0	225	2	T52011	ethylene responsiv
112	21	100.0	226	2	G69129	hypothetical prote
113	21	100.0	227	2	G89807	exotoxin 14 [impor
114	21	100.0	229	2	B90534	trna/rRNA methyltr
115	21	100.0	229	2	AG2390	hypothetical prote
116	21	100.0	232	1	HLMSE2	H-2 class II histo
117	21	100.0	233	2	I55654	MHC class II prote
118	21	100.0	234	2	PC4398	glutamine-tRNA lig
119	21	100.0	237	2	A21200	H-2 class II histo
120	21	100.0	237	2	T46121	hypothetical prote
121	21	100.0	241	2	I48657	I-E(b-beta) protei
122	21	100.0	245	2	B64007	hypothetical prote
123	21	100.0	246	2	T10127	transcription init
124	21	100.0	247	2	A81388	3-oxoacyl-[acyl-ca
125	21	100.0	248	2	D90986	hypothetical prote
126	21	100.0	248	2	G85831	hypothetical prote
127	21	100.0	248	2	F64971	hypothetical prote

128	21	100.0	248	2	AD0769	probable glycosylt
129	21	100.0	249	2	E75193	hypothetical prote
130	21	100.0	249	2	AF3225	hypothetical prote
131	21	100.0	250	2	F81961	probable hydroxyac
132	21	100.0	250	2	B86232	hypothetical prote
133	21	100.0	252	2	A95071	amino acid ABC tra
134	21	100.0	252	2	H81017	hydroxyacylglutath
135	21	100.0	252	2	AB1748	acetylerase hom
136	21	100.0	252	2	AI1378	acetylerase hom
137	21	100.0	255	2	S04668	hypothetical prote
138	21	100.0	256	2	AI3416	hypothetical membr
139	21	100.0	257	1	S40985	H+-transporting tw
140	21	100.0	258	2	D75520	conserved hypothet
141	21	100.0	258	2	G83417	probable hydroxyac
142	21	100.0	258	2	S77158	hydrogenase access
143	21	100.0	260	2	T15814	hypothetical prote
144	21	100.0	261	2	D70112	glutamate racemase
145	21	100.0	261	2	A69142	hypothetical prote
146	21	100.0	261	2	C84584	probable cAMP-depe
147	21	100.0	262	2	H45734	3'-aminoglycoside
148	21	100.0	262	2	C86776	kanamycin kinase (
149	21	100.0	262	2	T06600	acetyl-CoA carboxy
150	21	100.0	263	2	T16329	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

B60583

glycoprotein hormones alpha chain - bullfrog (fragment)

N;Alternate names: follitropin alpha chain; gonadotropin alpha chain; lutropin alpha chain; thyrotropin alpha chain

C;Species: Rana catesbeiana (bullfrog)

C;Date: 17-Apr-1993 #sequence\_revision 30-Sep-1993 #text\_change 17-Mar-1999

C;Accession: B60583

R;Bergert, E.R.; Madden, B.; McCormick, D.J.; Papkoff, H.; Ryan, R.J.

Endocrinology 127, 2985-2989, 1990

A;Title: The antigenic structure of the human glycoprotein hormone alpha-subunit: Cross-species comparisons.

A;Reference number: A60583; MUID:91065269; PMID:1701134

A;Accession: B60583

A;Molecule type: protein

A;Residues: 1-28 <BER>

C;Superfamily: glycoprotein hormones alpha chain

C;Keywords: glycoprotein; hormone

Query Match 100.0%; Score 21; DB 2; Length 28;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4

||||

Db 19 NLRF 22

##### RESULT 2

T44391  
 ribosomal protein L16 [imported] - *Bacillus halodurans*  
 C;Species: *Bacillus halodurans*  
 C;Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
 C;Accession: T44391  
 R;Takami, H.; Takaki, Y.; Nakasone, K.; Hiramata, C.; Inoue, A.; Horikoshi, K.  
 Biosci. Biotechnol. Biochem. 63, 452-455, 1999  
 A;Title: Sequence analysis of a 32-kb region including the major ribosomal  
 protein gene clusters from alkaliphilic *Bacillus* sp. strain C-125.  
 A;Reference number: Z22756; MUID:99209008; PMID:10192928  
 A;Accession: T44391  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-43 <TAK>  
 A;Cross-references: EMBL:AB017508; NID:g4512395; PIDN:BAA75279.1; PID:g4512412  
 A;Experimental source: strain C-125  
 C;Genetics:  
 A;Gene: rpmC  
 C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 43;  
 Best Local Similarity 100.0%; Pred. No. 38;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
 ||||  
 Db 3 NLR 6

# RESULT 3

S11603  
 hypothetical protein 2 (16S rRNA 3' region) - *Thermoplasma acidophilum*  
 (fragment)  
 C;Species: *Thermoplasma acidophilum*  
 C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 18-Jun-1993  
 C;Accession: S11603  
 R;Ree, H.K.; Cao, K.; Thurlow, D.L.; Zimmermann, R.A.  
 Can. J. Microbiol. 35, 124-133, 1989  
 A;Title: The structure and organization of the 16S ribosomal RNA gene from the  
 archaeobacterium *Thermoplasma acidophilum*.  
 A;Reference number: S11602; MUID:89248669; PMID:2470478  
 A;Accession: S11603  
 A;Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 1-50 <REE>

Query Match 100.0%; Score 21; DB 2; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
 ||||  
 Db 27 NLR 30

# RESULT 4

T18774



hypothetical protein B0513.3 - *Caenorhabditis elegans*  
 C;Species: *Caenorhabditis elegans*  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Aug-2000  
 C;Accession: T18774  
 R;Matthews, L.  
 submitted to the EMBL Data Library, November 1996  
 A;Reference number: Z19020  
 A;Accession: T18774  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-62 <WIL>  
 A;Cross-references: EMBL:Z82256; PIDN:CAB05115.1; GSPDB:GN00022; CESP:B0513.3  
 A;Experimental source: clone B0513  
 C;Genetics:  
 A;Gene: CESP:B0513.3  
 A;Map position: 4  
 A;Introns: 13/1  
 C;Superfamily: rat ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 62;  
 Best Local Similarity 100.0%; Pred. No. 57;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLR 4  
 ||||  
 Db 42 NLR 45

# RESULT 5

A87405  
 ribosomal protein L29 [imported] - *Caulobacter crescentus*  
 C;Species: *Caulobacter crescentus*  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
 C;Accession: A87405  
 R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.;  
 Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.;  
 Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.;  
 Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.;  
 Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf,  
 A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.;  
 Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A;Title: Complete Genome Sequence of *Caulobacter crescentus*.  
 A;Reference number: A87249; MUID:21173698; PMID:11259647  
 A;Accession: A87405  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-63 <STO>  
 A;Cross-references: GB:AE005673; NID:g13422587; PIDN:AAK23237.1; GSPDB:GN00148  
 C;Genetics:  
 A;Gene: CC1256  
 C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 63;  
 Best Local Similarity 100.0%; Pred. No. 58;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 27 NLR 30

RESULT 6

AG1778

ribosomal protein L29 [imported] - *Listeria innocua* (strain Clip11262)

C;Species: *Listeria innocua*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 17-May-2002

C;Accession: AG1778

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsiek, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.

A;Title: Comparative genomics of *Listeria* species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AG1778

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-63 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97999.1; PID:g16415309; GSPDB:GN00178

A;Experimental source: strain Clip11262

C;Genetics:

A;Gene: rpmC

C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 27 NLR 30

RESULT 7

AH1402

ribosomal protein L29 [imported] - *Listeria monocytogenes* (strain EGD-e)

C;Species: *Listeria monocytogenes*

C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 17-May-2002

C;Accession: AH1402

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, H.; Brandt, P.; Chakraborty, T.; Charbit, A.; Chetouani, F.; Couve, E.; de Daruvar, A.; Dehoux, P.; Domann, E.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Garcia-Del Portillo, F.; Garrido, P.; Gautier, L.; Goebel, W.; Gomez-Lopez, N.; Hain, T.; Hauf, J.; Jackson, D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mata Vicente, J.; Ng, E.; Nordsieck, G.; Novella, S.; de Pablos, B.; Perez-Diaz, J.C.; Remmel, B.; Rose, M.; Rusniok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, J.; Cossart, P.  
A;Title: Comparative genomics of *Listeria* species.  
A;Reference number: AB1077; MUID:21537279; PMID:11679669  
A;Accession: AH1402  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-63 <GLA>  
A;Cross-references: GB:NC\_003210; PIDN:CAD00702.1; PID:g16412112; GSPDB:GN00177  
A;Experimental source: strain EGD-e  
C;Genetics:  
A;Gene: rpmC  
C;Superfamily: Escherichia coli ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 63;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
|||  
Db 27 NLRf 30

#### RESULT 8

H90959

hypothetical protein ECs2648 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C;Accession: H90959

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: H90959

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-64 <HAY>

A;Cross-references: GB:BA000007; PIDN:BAB36071.1; PID:g13362116; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: ECs2648

Query Match 100.0%; Score 21; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
|||  
Db 14 NLRf 17

RESULT 9

R5BS29

ribosomal protein L29 - *Bacillus stearothermophilus*

C;Species: *Bacillus stearothermophilus*

C;Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 07-May-1999

C;Accession: A02825; S59065

R;Kimura, M.; Kimura, J.; Ashman, K.

Eur. J. Biochem. 150, 491-497, 1985

A;Title: The complete primary structure of ribosomal proteins L1, L14, L15, L23, L24, and L29 from *Bacillus stearothermophilus*.

A;Reference number: A91149; MUID:85257681; PMID:4018095

A;Accession: A02825

A;Molecule type: protein

A;Residues: 1-66 <KIM>

R;Urlaub, H.; Kruff, V.; Bischof, O.; Mueller, E.C.; Wittmann-Liebold, B.

EMBO J. 14, 4578-4588, 1995

A;Title: Protein-rRNA binding features and their structural and functional implications in ribosomes as determined by cross-linking studies.

A;Reference number: S59051; MUID:96003638; PMID:7556101

A;Accession: S59065

A;Molecule type: protein

A;Residues: 3-17;55-66 <URL>

C;Superfamily: *Escherichia coli* ribosomal protein L29

C;Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 21; DB 1; Length 66;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 27 NLRF 30

RESULT 10

R5BS2L

ribosomal protein L29 - *Bacillus subtilis*

C;Species: *Bacillus subtilis*

C;Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Jun-2000

C;Accession: S05990; E69697

R;Henkin, T.M.; Moon, S.H.; Mattheakis, L.C.; Nomura, M.

Nucleic Acids Res. 17, 7469-7486, 1989

A;Title: Cloning and analysis of the *spc* ribosomal protein operon of *Bacillus subtilis*: comparison with the *spc* operon of *Escherichia coli*.

A;Reference number: S05989; MUID:90016806; PMID:2508062

A;Accession: S05990

A;Molecule type: DNA

A;Residues: 1-66 <HEN>

A;Cross-references: EMBL:X15664; NID:g40146; PIDN:CAA33699.1; PID:g40148

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo,

V.; Bertero, M.G.; Bessieres, P.; Bolotin, A.; Borchert, S.; Boriss, R.;

Boursier, L.; Brans, A.; Braun, M.; Brignell, S.C.; Bron, S.; Brouillet, S.;

Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, S.K.; Codani,

J.J.; Connerton, I.F.; Cummings, N.J.; Daniel, R.A.; Denizot, F.; Devine, K.M.;

Duesterhoeft, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.;

Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim, S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.; Guiseppi, G.; Guy, B.J.; Haga, K.; Haiech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.; Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi, Y.; Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, S.

A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno, M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, D.; Porwollik, S.; Prescott, A.M.; Presecan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.; Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, E.

A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seror, S.J.; Serror, P.; Shin, B.S.; Soldo, B.; Sorokin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.; Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.; Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.; Wambutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: E69697

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-66 <KUN>

A;Cross-references: GB:Z99104; GB:AL009126; NID:g2632267; PIDN:CAB11900.1; PID:g2632391

A;Experimental source: strain 168

C;Genetics:

A;Gene: rpmC

C;Superfamily: *Escherichia coli* ribosomal protein L29

C;Keywords: protein biosynthesis; ribosome

Query Match 100.0%; Score 21; DB 1; Length 66;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLR 4  
|||  
Db 27 NLR 30

RESULT 11

AG3347

LSU ribosomal protein L29P [imported] - *Brucella melitensis* (strain 16M)

C;Species: *Brucella melitensis*

C;Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 17-May-2002

C;Accession: AG3347

R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova, N.; Anderson, I.; Bhattacharyya, A.; Lykidis, A.; Reznik, G.; Jablonski, L.; Larsen, N.; D'Souza, M.; Bernal, A.; Mazur, M.; Goltsman, E.;

Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letesson, J.J.; Haselkorn, R.; Kyrpides, N.; Overbeek, R.  
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
 A;Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*.  
 A;Reference number: AD3252; PMID:11756688  
 A;Accession: AG3347  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-66 <KUR>  
 A;Cross-references: GB:AE008917; PIDN:AAL51946.1; PID:g17982704; GSPDB:GN00190  
 A;Experimental source: strain 16M  
 C;Genetics:  
 A;Gene: BMEI0765  
 A;Map position: I  
 C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 66;  
 Best Local Similarity 100.0%; Pred. No. 61;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
 ||||  
 Db 27 NLRF 30

# RESULT 12

AH2814  
 50S ribosomal protein L29 [imported] - *Agrobacterium tumefaciens* (strain C58, Dupont)  
 C;Species: *Agrobacterium tumefaciens*  
 C;Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 18-Nov-2002  
 C;Accession: AH2814  
 R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.; Kitajima, J.P.; Okura, V.K.; Almeida Jr., N.F.; Zhou, Y.; Bovee Sr., D.; Chapman, P.; Clendenning, J.; Deatherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kuttyavin, T.; Levy, R.; Li, M.; McClelland, E.; Palmieri, A.; Raymond, C.; Rouse, G.; Saenphimmachak, C.; Wu, Z.; Gordon, D.; Eisen, J.A.; Paulsen, I.; Karp, P.; Romero, P.; Zhang, S.  
 Science 294, 2317-2323, 2001  
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, B.; Liao, L.; Kim, S.; Hendrick, C.; Zhao, Z.; Dolan, M.; Tingey, S.V.; Tomb, J.; Gordon, M.P.; Olson, M.V.; Nester, E.W.  
 A;Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.  
 A;Reference number: AB2577; MUID:21608550; PMID:11743193  
 A;Accession: AH2814  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-66 <KUR>  
 A;Cross-references: GB:AE008688; PIDN:AAL42934.1; PID:g17740391; GSPDB:GN00186  
 A;Experimental source: strain C58 (Dupont)  
 C;Genetics:  
 A;Gene: rpmC  
 A;Map position: circular chromosome  
 C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 27 NLR 30

RESULT 13

C97593

50S ribosomal protein L29 [imported] - *Agrobacterium tumefaciens* (strain C58, Cereon)

C;Species: *Agrobacterium tumefaciens*

C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 18-Nov-2002

C;Accession: C97593

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman, B.S.; Cao, Y.; Askenazi, M.; Halling, C.; Mullin, L.; Houmiel, K.; Gordon, J.; Vaudin, M.; Iartchouk, O.; Epp, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Flanagan, C.; Crowell, C.; Gurson, J.; Lomo, C.; Sear, C.; Strub, G.; Cielo, C.; Slater, S.

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium tumefaciens* C58.

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: C97593

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-66 <KUR>

A;Cross-references: GB:AE007869; PIDN:AAK87700.1; PID:g15157061; GSPDB:GN00169

C;Genetics:

A;Gene: AGR\_C\_3543

A;Map position: circular chromosome

C;Superfamily: *Escherichia coli* ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 66;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 27 NLR 30

RESULT 14

F83667

ribosomal protein L29 rpmC [imported] - *Bacillus halodurans* (strain C-125)

C;Species: *Bacillus halodurans*

C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 17-May-2002

C;Accession: F83667

R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hirama, C.; Nakamura, Y.; Ogasawara, N.; Kuhara, S.; Horikoshi, K.  
Nucleic Acids Res. 28, 4317-4331, 2000

A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and genomic sequence comparison with *Bacillus subtilis*.

A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: F83667

A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-67 <STO>  
A;Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BAB03861.1; GSPDB:GN00137  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: rpmC  
C;Superfamily: Escherichia coli ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 67;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 27 NLR 30

#### RESULT 15

E90021  
50S ribosomal protein L29 [imported] - Staphylococcus aureus (strain N315)  
C;Species: Staphylococcus aureus  
C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 17-May-2002  
C;Accession: E90021  
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Aoki, K.; Nagai, Y.; Lian, J.; Ito, T.; Kanamori, M.; Matsumaru, H.; Maruyama, A.; Murakami, H.; Hosoyama, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Hirakawa, H.; Kuhara, S.; Goto, S.; Yabuzaki, J.; Kanehisa, M.; Yamashita, A.; Oshima, K.; Furuya, K.; Yoshino, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.  
Lancet 357, 1225-1240, 2001  
A;Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.  
A;Reference number: A89758; MUID:21311952; PMID:11418146  
A;Accession: E90021  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-69 <KUR>  
A;Cross-references: GB:BA000018; PID:g13702042; PIDN:BAB43334.1; GSPDB:GN00149  
A;Experimental source: strain N315  
C;Genetics:  
A;Gene: rpmC  
C;Superfamily: Escherichia coli ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 69;  
Best Local Similarity 100.0%; Pred. No. 64;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 27 NLR 30

#### RESULT 16

A71363  
hypothetical protein TP0132 - syphilis spirochete  
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)



C;Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 05-Nov-1999  
 C;Accession: A71363  
 R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinn, M.; Hickey, E.K.; Clayton, R.; Ketchum, K.A.; Sodergren, E.; Hardham, J.M.; McLeod, M.P.; Salzberg, S.; Peterson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDonald, L.; Artiach, P.; Bowman, C.; Cotton, M.D.; Fujii, C.; Garland, S.; Hatch, B.; Horst, K.; Roberts, K.; Watthey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.  
 Science 281, 375-388, 1998  
 A;Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete.  
 A;Reference number: A71250; MUID:98332770; PMID:9665876  
 A;Accession: A71363  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-69 <COL>  
 A;Cross-references: GB:AE001198; GB:AE000520; NID:g3322390; PIDN:AAC65123.1; PID:g3322398  
 A;Experimental source: strain Nichols  
 C;Genetics:  
 A;Gene: TP0132

Query Match 100.0%; Score 21; DB 2; Length 69;  
 Best Local Similarity 100.0%; Pred. No. 64;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
 ||||  
 Db 9 NLRF 12

# RESULT 17

A71671  
 ribosomal protein L29 - *Rickettsia prowazekii*  
 C;Species: *Rickettsia prowazekii*  
 C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000  
 C;Accession: A71671  
 R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.C.M.; Podowski, R.M.; Naeslund, A.K.; Eriksson, A.S.; Winkler, H.H.; Kurland, C.G.  
 Nature 396, 133-140, 1998  
 A;Title: The genome sequence of *Rickettsia prowazekii* and the origin of mitochondria.  
 A;Reference number: A71630; MUID:99039499; PMID:9823893  
 A;Accession: A71671  
 A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-71 <AND>  
 A;Cross-references: GB:AJ235272; GB:AJ235269; NID:g3861033; PIDN:CAA15091.1; PID:e1342935; PID:g3861191; GSPDB:GN00081  
 A;Experimental source: strain Madrid E  
 C;Genetics:  
 A;Gene: rpmC; RP651

Query Match 100.0%; Score 21; DB 2; Length 71;  
 Best Local Similarity 100.0%; Pred. No. 66;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy            1 NLRF 4  
               ||||  
 Db           32 NLRF 35

RESULT 18

F97824

50S ribosomal protein L29 [imported] - Rickettsia conorii (strain Malish 7)

C;Species: Rickettsia conorii

C;Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 30-Sep-2001

C;Accession: F97824

R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.;  
 Samson, D.; Roux, V.; Cossart, P.; Weissenbach, J.; Claverie, J.M.; Raoult, D.  
 Science 293, 2093-2098, 2001

A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia  
 prowazekii.

A;Reference number: A97700; MUID:21442074; PMID:11557893

A;Accession: F97824

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-71 <KUR>

A;Cross-references: GB:AE006914; PIDN:AAL03536.1; PID:g15620113; GSPDB:GN00173

C;Genetics:

A;Gene: rpmC

Query Match                    100.0%;   Score 21;   DB 2;   Length 71;  
 Best Local Similarity       100.0%;   Pred. No. 66;  
 Matches       4;   Conservative       0;   Mismatches       0;   Indels       0;   Gaps       0;

Qy            1 NLRF 4  
               ||||  
 Db           32 NLRF 35

RESULT 19

H70642

probable ribosomal protein L29 rpmC - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C;Accession: H70642

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.;  
 Gordon, S.V.; Eiglmeier, K.; Gas, S.; Barry III, C.E.; Tekaia, F.; Badcock, K.;  
 Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;  
 Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.;  
 Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail,  
 M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.;  
 Squares, S.

Nature 393, 537-544, 1998

A;Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete  
 genome sequence.

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: H70642

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-77 <COL>

A;Cross-references: GB:Z84395; GB:AL123456; NID:g3261698; PIDN:CAB06433.1;  
PID:g1806177  
A;Experimental source: strain H37Rv  
C;Genetics:  
A;Gene: rpmC  
C;Superfamily: Escherichia coli ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 77;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLRF 4  
|||  
Db 31 NLRF 34

RESULT 20

T17858  
hypothetical protein a358R - Chlorella virus PBCV-1  
C;Species: Chlorella virus PBCV-1  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C;Accession: T17858  
R;Graves, M.V.; Van Etten, J.L.  
submitted to the EMBL Data Library, May 1999  
A;Reference number: Z18806  
A;Accession: T17858  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-77 <GRA>  
A;Cross-references: EMBL:U42580; NID:g4028896; PIDN:AAC96726.1  
A;Experimental source: specific host Chlorella strain NC64A  
C;Genetics:  
A;Note: a358R

Query Match 100.0%; Score 21; DB 2; Length 77;  
Best Local Similarity 100.0%; Pred. No. 72;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLRF 4  
|||  
Db 55 NLRF 58

RESULT 21

B90806  
hypothetical protein ECs1418 [imported] - Escherichia coli (strain O157:H7,  
substrain RIMD 0509952)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
C;Accession: B90806  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.;  
Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida,  
T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara,  
S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7  
and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: B90806  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-79 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BAB34841.1; PID:g13360878; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs1418

Query Match 100.0%; Score 21; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 43 NLRF 46

#### RESULT 22

C69036

hypothetical protein MTH1266 - Methanobacterium thermoautotrophicum (strain Delta H)

C;Species: Methanobacterium thermoautotrophicum

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999

C;Accession: C69036

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.; Bashirzadeh, R.; Blakely, D.; Cook, R.; Gilbert, K.; Harrison, D.; Hoang, L.; Keagle, P.; Lumm, W.; Pothier, B.; Qiu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.; Caruso, A.; Bush, D.; Safer, H.; Patwell, D.; Prabhakar, S.; McDougall, S.; Shimer, G.; Goyal, A.; Pietrokovski, S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: functional analysis and comparative genomics.

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: C69036

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-79 <MTH>

A;Cross-references: GB:AE000893; GB:AE000666; NID:g2622375; PIDN:AAB85755.1; PID:g2622380

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: MTH1266

Query Match 100.0%; Score 21; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 74;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 40 NLRF 43

#### RESULT 23

T45372  
 ribosomal protein L29 [imported] - Mycobacterium leprae  
 C;Species: Mycobacterium leprae  
 C;Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
 C;Accession: T45372  
 R;Cole, S.T.; Flesselles, B.; Honore, N.  
 submitted to the EMBL Data Library, August 1997  
 A;Reference number: Z22966  
 A;Accession: T45372  
 A;Status: preliminary; translated from GB/EMBL/DDBJ  
 A;Molecule type: DNA  
 A;Residues: 1-80 <COL>  
 A;Cross-references: EMBL:Z98756; PIDN:CAB11442.1  
 A;Experimental source: cosmid B2492  
 C;Genetics:  
 A;Note: rpmC  
 C;Superfamily: Escherichia coli ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 75;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
 ||||  
 Db 31 NLRF 34

#### RESULT 24

A87141  
 50S ribosomal protein L29 [imported] - Mycobacterium leprae  
 C;Species: Mycobacterium leprae  
 C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 17-May-2002  
 C;Accession: A87141  
 R;Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Honore, N.; Ganier, T.; Churcher, C.; Harris, D.; Mungall, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Lacroix, C.; Maclean, J.; Moule, S.; Murphy, L.; Oliver, K.; Quail, M.A.; Rajandream, M.A.; Rutherford, K.M.  
 Nature 409, 1007-1011, 2001  
 A;Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squares, S.; Stevens, K.; Taylor, K.; Whitehead, S.; Woodward, J.R.; Barrell, B.G.  
 A;Title: Massive gene decay in the leprosy bacillus.  
 A;Reference number: A86909; MUID:21128732; PMID:11234002  
 A;Accession: A87141  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-81 <STO>  
 A;Cross-references: GB:AL450380; NID:g13093547; PIDN:CAC30809.1; GSPDB:GN00147  
 C;Genetics:  
 A;Gene: rpmC  
 C;Superfamily: Escherichia coli ribosomal protein L29

Query Match 100.0%; Score 21; DB 2; Length 81;  
 Best Local Similarity 100.0%; Pred. No. 76;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy. 1 NLRF 4  
 ||||  
Db 32 NLRF 35

RESULT 25

S71166

RNA-directed DNA polymerase (EC 2.7.7.49) - Arabidopsis thaliana retrotransposon  
3 (fragment)

N;Alternate names: reverse transcriptase

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 28-Oct-1996 #sequence\_revision 27-Feb-1997 #text\_change 20-Jun-2000

C;Accession: S71166

R;Hirochika, H.; Hirochika, R.

submitted to the EMBL Data Library, August 1992

A;Description: Retrotransposons are ubiquitous in plants.

A;Reference number: S71166

A;Accession: S71166

A;Molecule type: DNA

A;Residues: 1-82 <HIR>

A;Cross-references: EMBL:D12843; NID:g217866; PID:g217867

C;Genetics:

A;Mobile element: retrotransposon 3

C;Superfamily: retrovirus-related polyprotein

C;Keywords: nucleotidyltransferase; reverse transcriptase

Query Match 100.0%; Score 21; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
 ||||  
Db 46 NLRF 49

RESULT 26

E86562

natural UGA frame-shift [imported] - Chlamydophila pneumoniae (strain J138)

C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 02-Mar-2001

C;Accession: E86562

R;Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.;

Shiba, T.; Ishii, K.; Hattori, M.; Kuhara, S.; Nakazawa, T.

Nucleic Acids Res. 28, 2311-2314, 2000

A;Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.

A;Reference number: A86491; MUID:20330349; PMID:10871362

A;Accession: E86562

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-82 <STO>

A;Cross-references: GB:BA000008; NID:g8978948; PIDN:BAA98783.1; GSPDB:GN00142

A;Experimental source: strain J138

C;Genetics:

A;Gene: CPj0576\_2

Query Match 100.0%; Score 21; DB 2; Length 82;

Best Local Similarity 100.0%; Pred. No. 77;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 75 NLRF 78

RESULT 27

T15818

hypothetical protein C52B11.4 - *Caenorhabditis elegans*

C;Species: *Caenorhabditis elegans*

C;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C;Accession: T15818

R;Martin, J.

submitted to the EMBL Data Library, November 1995

A;Description: The sequence of *C. elegans* cosmid C52B11.

A;Reference number: Z18411

A;Accession: T15818

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-82 <MAR>

A;Cross-references: EMBL:U41276; NID:g1086884; PID:g1086889; PIDN:AAA82471.1;

CESP:C52B11.4

A;Experimental source: strain Bristol N2

C;Genetics:

A;Gene: CESP:C52B11.4

A;Introns: 6/1; 34/2

Query Match 100.0%; Score 21; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 38 NLRF 41

RESULT 28

T07204

hypothetical protein 86b - *Chlorella vulgaris* chloroplast

C;Species: chloroplast *Chlorella vulgaris*

C;Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 21-Jul-2000

C;Accession: T07204

R;Wakasugi, T.; Nagai, T.; Kapoor, M.; Sugita, M.; Ito, M.; Ito, S.; Tsudzuki, J.; Nakashima, K.; Tsudzuki, T.; Suzuki, Y.; Hamada, A.; Ohta, T.; Inamura, A.; Yoshinaga, K.; Sugiura, M.

Proc. Natl. Acad. Sci. U.S.A. 94, 5967-5972, 1997

A;Title: Complete nucleotide sequence of the chloroplast genome from the green alga *Chlorella vulgaris*: the existence of genes possibly involved in chloroplast division.

A;Reference number: Z15985; MUID:97303241; PMID:9159184

A;Accession: T07204

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-86 <WAK>

A;Cross-references: EMBL:AB001684; NID:g2224352; PIDN:BAA57851.1; PID:g2224367

C;Genetics:  
A;Genome: chloroplast  
C;Keywords: chloroplast

Query Match 100.0%; Score 21; DB 2; Length 86;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 12 NLRF 15

RESULT 29

S38681

major histocompatibility complex class II beta chain - northern lesser bushbaby  
(fragment)

N;Alternate names: MHC class II beta chain

C;Species: Galago senegalensis (northern lesser bushbaby)

C;Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 19-May-2000

C;Accession: S38681

R;Figuroa, F.; O'Huigin, C.; Tichy, H.; Klein, J.

submitted to the EMBL Data Library, November 1993

A;Description: The origin of the primate Mhc-DRB genes and allelic lineages as  
deduced from the study of prosimians.

A;Reference number: S38676

A;Accession: S38681

A;Molecule type: DNA

A;Residues: 1-87 <FIG>

A;Cross-references: EMBL:Z27151

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 21; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 32 NLRF 35

RESULT 30

S38682

major histocompatibility complex class II beta chain - northern lesser bushbaby  
(fragment)

N;Alternate names: MHC class II beta chain

C;Species: Galago senegalensis (northern lesser bushbaby)

C;Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 19-May-2000

C;Accession: S38682

R;Figuroa, F.; O'Huigin, C.; Tichy, H.; Klein, J.

submitted to the EMBL Data Library, November 1993

A;Description: The origin of the primate Mhc-DRB genes and allelic lineages as  
deduced from the study of prosimians.

A;Reference number: S38676

A;Accession: S38682

A;Molecule type: DNA

A;Residues: 1-87 <FIG>



A;Cross-references: EMBL:Z27152

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 21; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLRF 4  
||||  
Db 32 NLRF 35

RESULT 31

S38685

major histocompatibility complex class II beta chain - northern lesser bushbaby (fragment)

N;Alternate names: MHC class II beta chain

C;Species: Galago senegalensis (northern lesser bushbaby)

C;Date: 19-Mar-1997 #sequence\_revision 06-Jun-1997 #text\_change 19-May-2000

C;Accession: S38685

R;Figuroa, F.; O'Huigin, C.; Tichy, H.; Klein, J.

submitted to the EMBL Data Library, November 1993

A;Description: The origin of the primate Mhc-DRB genes and allelic lineages as deduced from the study of prosimians.

A;Reference number: S38676

A;Accession: S38685

A;Molecule type: DNA

A;Residues: 1-87 <FIG>

A;Cross-references: EMBL:Z27155

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 21; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLRF 4  
||||  
Db 32 NLRF 35

RESULT 32

S03117

class II histocompatibility antigen H-2-I-E beta chain - mouse (fragment)

C;Species: Mus musculus (house mouse)

A;Variety: strain B10.S

C;Date: 07-Sep-1990 #sequence\_revision 01-Nov-1996 #text\_change 23-Jul-1999

C;Accession: S03117

R;Stone, J.; Prey, R.; Todd, J.A.; McDevitt, H.O.

submitted to the EMBL Data Library, March 1988

A;Description: Nucleotide sequences of the murine Ia-associated invariant chain (Ii) and I-E (H2S, beta) chain expressible cDNA clones.

A;Reference number: S03117

A;Accession: S03117

A;Molecule type: mRNA

A;Residues: 1-87 <STO>

A;Cross-references: EMBL:X07239; NID:g51528; PIDN:CAA30228.1; PID:g51529

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

C;Keywords: heterodimer

Query Match 100.0%; Score 21; DB 2; Length 87;  
Best Local Similarity 100.0%; Pred. No. 82;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
|||  
Db 64 NLRf 67

RESULT 33

D47759

retrovirus-related reverse transcriptase homolog - common sunflower  
retrotransposon copia-like (fragment)  
C;Species: Helianthus annuus (common sunflower)  
C;Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 21-Jul-2000  
C;Accession: D47759  
R;Voytas, D.F.; Cummings, M.P.; Koniczny, A.; Ausubel, F.M.; Rodermel, S.R.  
Proc. Natl. Acad. Sci. U.S.A. 89, 7124-7128, 1992  
A;Title: copia-like retrotransposons are ubiquitous among plants.  
A;Reference number: A46200; MUID:92357784; PMID:1379734  
A;Accession: D47759  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-88 <VOY>  
A;Cross-references: GB:M94494; NID:g439421; PIDN:AAA33373.1; PID:g168235  
A;Note: sequence extracted from NCBI backbone (NCBIP:111888)  
C;Superfamily: retrovirus-related polyprotein

Query Match 100.0%; Score 21; DB 2; Length 88;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
|||  
Db 48 NLRf 51

RESULT 34

S38676

MHC class II histocompatibility antigen HLA-DR-01 beta chain - Galago moholi  
(fragment)  
C;Species: Galago moholi  
C;Date: 25-Dec-1994 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
C;Accession: S38676  
R;Figueroa, F.; O'Huigin, C.; Tichy, H.; Klein, J.  
submitted to the EMBL Data Library, November 1993  
A;Description: The origin of the primate Mhc-DRB genes and allelic lineages as  
deduced from the study of prosimians.  
A;Reference number: S38676  
A;Accession: S38676  
A;Molecule type: DNA  
A;Residues: 1-89 <FIG>  
A;Cross-references: EMBL:Z27130; NID:g415801; PIDN:CAA81657.1; PID:g1132545  
C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 21; DB 2; Length 89;  
Best Local Similarity 100.0%; Pred. No. 84;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
||||  
Db 32 NLRf 35

RESULT 35

G71630

integration host factor alpha-chain (himA) RP708 - Rickettsia prowazekii

C;Species: Rickettsia prowazekii

C;Date: 21-Nov-1998 #sequence\_revision 21-Nov-1998 #text\_change 03-Nov-2000

C;Accession: G71630

R;Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alsmark, U.C.M.; Podowski, R.M.; Naeslund, A.K.; Eriksson, A.S.; Winkler, H.H.; Kurland, C.G.

Nature 396, 133-140, 1998

A;Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A;Reference number: A71630; MUID:99039499; PMID:9823893

A;Accession: G71630

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-95 <AND>

A;Cross-references: GB:AJ235273; GB:AJ235269; NID:g3861237; PIDN:CAA15143.1; PID:g3861244; GSPDB:GN00081

A;Experimental source: strain Madrid E

C;Genetics:

A;Gene: himA; RP708

C;Superfamily: bacterial DNA-binding protein

Query Match 100.0%; Score 21; DB 2; Length 95;  
Best Local Similarity 100.0%; Pred. No. 90;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
||||  
Db 77 NLRf 80

RESULT 36

E90894

hypothetical protein ECs2125 [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001

C;Accession: E90894

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Ohtsubo, E.; Nakayama, K.; Murata, T.; Tanaka, M.; Tobe, T.; Iida, T.; Takami, H.; Honda, T.; Sasakawa, C.; Ogasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: E90894  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-96 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BAB35548.1; PID:g13361591; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: ECs2125

Query Match 100.0%; Score 21; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
|||  
Db 35 NLRf 38

RESULT 37

D85723

hypothetical protein Z2187 [imported] - Escherichia coli (strain O157:H7, substrain EDL933)

C;Species: Escherichia coli

C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C;Accession: D85723

R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew, G.F.; Evans, P.S.; Gregor, J.; Kirkpatrick, H.A.; Posfai, G.; Hackett, J.; Klink, S.; Boutin, A.; Shao, Y.; Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca, J.; Anantharaman, T.S.; Lin, J.; Yen, G.; Schwartz, D.C.; Welch, R.A.; Blattner, F.R.

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A;Reference number: A85480; MUID:21074935; PMID:11206551

A;Accession: D85723

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-96 <STO>

A;Cross-references: GB:AE005174; NID:g12515148; PIDN:AAG56248.1; GSPDB:GN00145; UWGP:Z2187

A;Experimental source: strain O157:H7, substrain EDL933

C;Genetics:

A;Gene: Z2187

Query Match 100.0%; Score 21; DB 2; Length 96;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRf 4  
|||  
Db 35 NLRf 38

RESULT 38

AF0050

conserved hypothetical protein YPO0407 [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001  
 C;Accession: AF0050  
 R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.;  
 Prentice, M.B.; Sebaihia, M.; James, K.D.; Churcher, C.; Mungall, K.L.; Baker,  
 S.; Basham, D.; Bentley, S.D.; Brooks, K.; Cerdeno-Tarraga, A.M.; Chillingworth,  
 T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Feltwell, T.; Hamlin, N.;  
 Holroyd, S.; Jagels, K.; Leather, S.; Karlyshev, A.V.; Moule, S.; Oyston,  
 P.C.F.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
 Whitehead, S.; Barrell, B.G.  
 Nature 413, 523-527, 2001  
 A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
 A;Reference number: AB0001; MUID:21470413; PMID:11586360  
 A;Accession: AF0050  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-96 <KUR>  
 A;Cross-references: GB:AL590842; PIDN:CAC89265.1; PID:g15978504; GSPDB:GN00175  
 C;Genetics:  
 A;Gene: YPO0407

Query Match 100.0%; Score 21; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
 ||||  
 Db 35 NLR 38

# RESULT 39

A64906  
 hypothetical protein b1518 - *Escherichia coli* (strain K-12)  
 C;Species: *Escherichia coli*  
 C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 01-Mar-2002  
 C;Accession: A64906  
 R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.;  
 Riley, M.; Collado-Vides, J.; Glasner, J.D.; Rode, C.K.; Mayhew, G.F.; Gregor,  
 J.; Davis, N.W.; Kirkpatrick, H.A.; Goeden, M.A.; Rose, D.J.; Mau, B.; Shao, Y.  
 Science 277, 1453-1462, 1997  
 A;Title: The complete genome sequence of *Escherichia coli* K-12.  
 A;Reference number: A64720; MUID:97426617; PMID:9278503  
 A;Accession: A64906  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: DNA  
 A;Residues: 1-96 <BLAT>  
 A;Cross-references: GB:AE000249; GB:U00096; NID:g1787790; PIDN:AAC74591.1;  
 PID:g1787797; UWGP:b1518  
 A;Experimental source: strain K-12, substrain MG1655

Query Match 100.0%; Score 21; DB 2; Length 96;  
 Best Local Similarity 100.0%; Pred. No. 91;  
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
 ||||  
 Db 35 NLR 38

RESULT 40

S20287

lutropin alpha chain - bullfrog

C;Species: *Rana catesbeiana* (bullfrog)

C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 12-Dec-1997

C;Accession: S20287

R;Hayashi, H.; Hayashi, T.; Hanaoka, Y.

Eur. J. Biochem. 203, 185-191, 1992

A;Title: Amphibian lutropin and follitropin from the bullfrog *Rana catesbeiana*. Complete amino acid sequence of the alpha subunit.

A;Reference number: S20287; MUID:92111564; PMID:1730225

A;Accession: S20287

A;Molecule type: protein

A;Residues: 1-97 <HAY>

C;Superfamily: glycoprotein hormones alpha chain

C;Keywords: glycoprotein

F;57,83/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 100.0%; Score 21; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 19 NLR 22

RESULT 41

G87667

hypothetical protein CC3377 [imported] - *Caulobacter crescentus*

C;Species: *Caulobacter crescentus*

C;Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001

C;Accession: G87667

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.; Alley, M.; Ohta, N.; Maddock, J.R.; Potocka, I.; Nelson, W.C.; Newton, A.; Stephens, C.; Phadke, N.D.; Ely, B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolonay, J.F.; Smit, J.; Craven, M.; Khouri, H.; Shetty, J.; Berry, K.; Utterback, T.; Tran, K.; Wolf, A.; Vamathevan, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of *Caulobacter crescentus*.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: G87667

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-97 <STO>

A;Cross-references: GB:AE005673; NID:g13425083; PIDN:AAK25339.1; GSPDB:GN00148

C;Genetics:

A;Gene: CC3377

Query Match 100.0%; Score 21; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4

Db

||||  
49 NLRF 52

RESULT 42

AB0101

hypothetical protein YPO0820 [imported] - *Yersinia pestis* (strain CO92)

C;Species: *Yersinia pestis*

C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Nov-2001

C;Accession: AB0101

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.;  
Prentice, M.B.; Sebaihia, M.; James, K.D.; Churcher, C.; Mungall, K.L.; Baker,  
S.; Basham, D.; Bentley, S.D.; Brooks, K.; Cerdeno-Tarraga, A.M.; Chillingworth,  
T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Feltwell, T.; Hamlin, N.;  
Holroyd, S.; Jagels, K.; Leather, S.; Karlyshev, A.V.; Moule, S.; Oyston,  
P.C.F.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
Whitehead, S.; Barrell, B.G.

Nature 413, 523-527, 2001

A;Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AB0101

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-102 <KUR>

A;Cross-references: GB:AL590842; PIDN:CAC89669.1; PID:g15978896; GSPDB:GN00175

C;Genetics:

A;Gene: YPO0820

Query Match 100.0%; Score 21; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
||||  
Db 7 NLRF 10

RESULT 43

G72700

hypothetical protein APE1022 - *Aeropyrum pernix* (strain K1)

C;Species: *Aeropyrum pernix*

C;Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 20-Aug-1999

C;Accession: G72700

R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no,  
K.; Takahashi, M.; Sekine, M.; Baba, S.; Ankai, A.; Kosugi, H.; Hosoyama, A.;  
Fukui, S.; Nagai, Y.; Nishijima, K.; Nakazawa, H.; Takamiya, M.; Masuda, S.;  
Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.;  
Aoki, K.; Kubota, K.; Nakamura, Y.; Nomura, N.; Sako, Y.; Kikuchi, H.

DNA Res. 6, 83-101, 1999

A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon,  
*Aeropyrum pernix* K1.

A;Reference number: A72450; MUID:99310339; PMID:10382966

A;Accession: G72700

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-102 <KAW>

A;Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BAA80007.1; PID:d1043793;  
PID:g5104692  
A;Experimental source: strain K1  
C;Genetics:  
A;Gene: APE1022

Query Match 100.0%; Score 21; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 97;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NLR 4  
|||  
Db 93 NLR 96

#### RESULT 44

D95916

hypothetical protein SMB21023 [imported] - Sinorhizobium meliloti (strain 1021)  
megaplasmid pSymB

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001

C;Accession: D95916

R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernandez-Lucas, I.; Becker, A.; Cowie, A.; Gouzy, J.; Golding, B.; Puhler, A.

Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endosymbiont Sinorhizobium meliloti.

A;Reference number: A95842; MUID:21396508; PMID:11481431

A;Accession: D95916

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-104 <KUR>

A;Cross-references: GB:AL591985; PIDN:CAC48996.1; PID:g15140481; GSPDB:GN00167

A;Experimental source: strain 1021, megaplasmid pSymB

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.; Barnett, M.J.; Becker, A.; Boistard, P.; Bothe, G.; Boutry, M.; Bowser, L.; Buhrmester, J.; Cadieu, E.; Capela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; Gloux, S.; Godrie, T.; Goffeau, A.; Golding, B.; Gouzy, J.; Gurjal, M.; Hernandez-Lucas, I.; Hong, A.; Huizar, L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, V.; Masuy, D.; Palm, C.; Peck, M.C.; Pohl, T.M.; Portetelle, D.; Purnelle, B.; Ramsperger, U.; Surzycki, R.; Thebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.; Batut, J.

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: SMB21023

A;Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 99;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Qy            1 NLR 4  
             ||||  
Db            35 NLR 38

RESULT 45

T38761

hypothetical protein SPAC3H8.03 - fission yeast (*Schizosaccharomyces pombe*).

C;Species: *Schizosaccharomyces pombe*

C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C;Accession: T38761

R;Gentles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsh, S.V.  
submitted to the EMBL Data Library, January 1996

A;Reference number: Z21810

A;Accession: T38761

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-105 <GEN>

A;Cross-references: EMBL:Z69086; PIDN:CAA93160.1; GSPDB:GN00066; SPDB:SPAC3H8.03

A;Experimental source: strain 972h-; cosmid c3H8

C;Genetics:

A;Gene: SPDB:SPAC3H8.03

A;Map position: 1

A;Introns: 60/3

Query Match            100.0%;   Score 21;   DB 2;   Length 105;  
Best Local Similarity   100.0%;   Pred. No. 1e+02;  
Matches       4;   Conservative       0;   Mismatches       0;   Indels       0;   Gaps       0;

Qy            1 NLR 4  
             ||||  
Db            11 NLR 14

RESULT 46

S66828

hypothetical protein YOL131w - yeast (*Saccharomyces cerevisiae*)

N;Alternate names: hypothetical protein AOA108; hypothetical protein O0517

C;Species: *Saccharomyces cerevisiae*

C;Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 19-Apr-2002

C;Accession: S66828; S72037

R;Arino, J.; Casamayor, A.; Gamo, F.J.; Gancedo, C.; Lafuente, M.J.; Aldea, M.;  
Casas, C.; Herrero, E.

submitted to the Protein Sequence Database, July 1996

A;Reference number: S66814

A;Accession: S66828

A;Molecule type: DNA

A;Residues: 1-108 <ARI>

A;Cross-references: EMBL:Z74873; NID:g1420018; PID:g1420019; MIPS:YOL131w

A;Experimental source: strain S288C

R;Aldea, M.; Piedrafita, L.; Casas, C.; Casamayor, A.; Khalid, H.; Balcells, L.;  
Arino, J.; Herrero, E.

Yeast 12, 1053-1058, 1996

A;Title: Sequence analysis of a 12 801 bp fragment of the left arm of yeast  
chromosome XV containing a putative 6-phosphofructo-2-kinase gene, a gene for a  
possible glycopospholipid-anchored surface protein and six other open reading  
frames.

A;Reference number: S72030; MUID:97051593; PMID:8896270  
A;Accession: S72037  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-108 <ALD>  
A;Cross-references: EMBL:X95465; NID:g1628437; PIDN:CAA64739.1; PID:g1628445  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1996  
C;Genetics:  
A;Cross-references: SGD:S0005491  
A;Map position: 15L  
A;Note: YOL131w  
C;Superfamily: Saccharomyces hypothetical protein YOL131w

Query Match 100.0%; Score 21; DB 2; Length 108;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 27 NLRF 30

#### RESULT 47

AF0940

conserved hypothetical protein STY3791 [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)

C;Species: Salmonella enterica subsp. enterica serovar Typhi

A;Note: this species has also been called Salmonella typhi

C;Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C;Accession: AF0940

R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; Mungall, K.L.; Bentley, S.D.; Holden, M.T.G.; Sebaihia, M.; Baker, S.; Basham, D.; Brooks, K.; Chillingworth, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.; Feltwell, T.; Hamlin, N.; Hague, A.; Hien, T.T.; Holroyd, S.; Jagels, K.; Krogh, A.; Larsen, T.S.; Leather, S.; Moule, S.; O'Gaora, P.

Nature 413, 848-852, 2001

A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, B.G.

A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18.

A;Reference number: AB0502; MUID:21534947; PMID:11677608

A;Accession: AF0940

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-109 <PAR>

A;Cross-references: GB:AL513382; PIDN:CAD09544.1; PID:g16504660; GSPDB:GN00176

C;Genetics:

A;Gene: STY3791

Query Match 100.0%; Score 21; DB 2; Length 109;  
Best Local Similarity 100.0%; Pred. No. 1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||

## RESULT 48

S64730

protein secretion protein xcpT - *Pseudomonas putida* (fragment)C;Species: *Pseudomonas putida*

C;Date: 06-Dec-1996 #sequence\_revision 06-Dec-1996 #text\_change 08-Sep-1997

C;Accession: S64730; S47507

R;de Groot, A.; Krijger, J.J.; Filloux, A.; Tommassen, J.

Mol. Gen. Genet. 250, 491-504, 1996

A;Title: Characterization of type II protein secretion (xcp) genes in the plant growth-stimulating *Pseudomonas putida*, strain WCS358.

A;Reference number: S64724; MUID:96186881; PMID:8602167

A;Accession: S64730

A;Molecule type: DNA

A;Residues: 1-112 &lt;DEG&gt;

A;Cross-references: EMBL:X81085; NID:g531737; PID:g531743

C;Genetics:

A;Gene: xcpT

C;Superfamily: secretion protein xcpT

Query Match 100.0%; Score 21; DB 2; Length 112;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4

||||

Db 64 NLRF 67

## RESULT 49

C71221

hypothetical protein PH0030 - *Pyrococcus horikoshii*C;Species: *Pyrococcus horikoshii*

C;Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 21-Jul-2000

C;Accession: C71221

R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosugi, H.; Hosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.; Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.; Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.; Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H. DNA Res. 5, 55-76, 1998

A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, *Pyrococcus horikoshii* OT3.

A;Reference number: A71000; MUID:98344137; PMID:9679194

A;Accession: C71221

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-112 &lt;KAW&gt;

A;Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29098.1; PID:g3256415

A;Experimental source: strain OT3

A;Note: this accession replaces an interim accession for a sequence replaced by GenBank

C;Genetics:

A;Gene: PH0030

C;Superfamily: *Pyrococcus horikoshii* hypothetical protein PH0030

Query Match 100.0%; Score 21; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 26 NLR 29

RESULT 50

AF2549

hypothetical protein all7681 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120beta

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002

C;Accession: AF2549

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, M.; Ishikawa, A.; Kawashima, K.; Kimura, T.; Kishida, Y.; Kohara, M.; Matsumoto, M.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120.

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AF2549

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-112 <KUR>

A;Cross-references: GB:AP003602; PIDN:BAB77324.1; PID:g17134766; GSPDB:GN00181

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: all7681

A;Genome: plasmid

Query Match 100.0%; Score 21; DB 2; Length 112;  
Best Local Similarity 100.0%; Pred. No. 1.1e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLR 4  
|||  
Db 84 NLR 87

RESULT 51

G86654

transcription regulator [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C;Species: Lactococcus lactis subsp. lactis

C;Date: 23-Mar-2001 #sequence\_revision 23-Mar-2001 #text\_change 03-Aug-2001

C;Accession: G86654

R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Ehrlich, S.D.; Sorokin, A.

Genome Res. 11, 731-753, 2001

A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis IL1403.

A;Reference number: A86625; MUID:21235186; PMID:11337471  
A;Accession: G86654  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-120 <STO>  
A;Cross-references: GB:AE005176; PID:g12723097; PIDN:AAK04337.1; GSPDB:GN00146  
A;Experimental source: strain IL1403  
C;Genetics:  
A;Gene: rmeC

Query Match 100.0%; Score 21; DB 2; Length 120;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 17 NLRF 20

RESULT 52

A46561

MHC class II histocompatibility antigen H2-IE-k beta chain precursor - mouse (fragment)

C;Species: Mus musculus (house mouse)

C;Date: 03-Mar-1994 #sequence\_revision 03-Mar-1994 #text\_change 17-Mar-1999

C;Accession: A46561

R;King, L.B.; Sharma, S.; Corley, R.B.

J. Immunogenet. 15, 209-214, 1988

A;Title: Complete coding region sequence of E-beta(k) cDNA clones: lack of polymorphism in the NH-2-terminus between E-beta(k) and E-beta(b) molecules.

A;Reference number: A46561; MUID:89278714; PMID:3150763

A;Accession: A46561

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-121 <KIN>

A;Cross-references: GB:M36939; GB:M18579

A;Note: authors translated the codon GTG for residue 23 as Ile

C;Superfamily: class II histocompatibility antigen; immunoglobulin homology

Query Match 100.0%; Score 21; DB 2; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLRF 4  
|||  
Db 63 NLRF 66

RESULT 53

AI0204

chemotaxis protein CheY [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 02-Jun-2003

C;Accession: AI0204

R;Parkhill, J.; Wren, B.W.; Thomson,